

Variable-Metrik-Verfahren mit nichtquadratischen Näherungsfunktionen

INAUGURAL-DISSERTATION

zur Erlangung der Philosophischen Doktorwürde
vorgelegt der
Philosophischen Fakultät II
der
UNIVERSITÄT ZÜRICH

von

FORTUNAT MARTIN SCHMID

von St. Gallen

Begutachtet von den Herren
Prof. Dr. P. Kall
Prof. Dr. H. R. Schwarz

1986
Zürich ~~1985~~

Zentralstelle der Studentenschaft

Variable-Metrik-Verfahren mit nichtquadratischen Näherungsfunktionen

INAUGURAL-DISSERTATION

zur Erlangung der Philosophischen Doktorwürde
vorgelegt der
Philosophischen Fakultät II
der
UNIVERSITÄT ZÜRICH

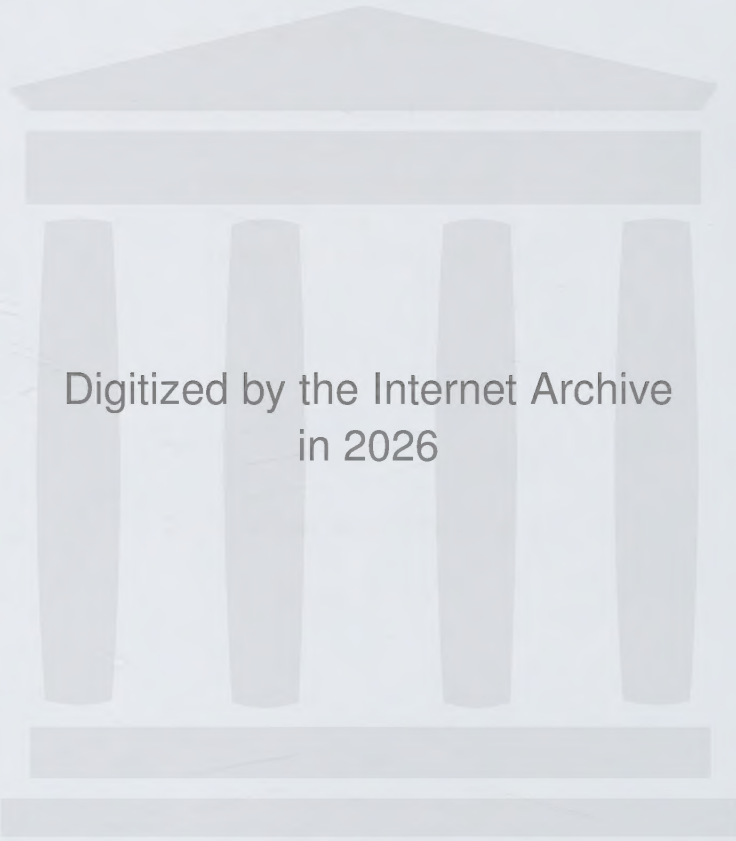
von

FORTUNAT MARTIN SCHMID

von St. Gallen

Begutachtet von den Herren
Prof. Dr. P. Kall
Prof. Dr. H. R. Schwarz

Zürich 1985
Zentralstelle der Studentenschaft



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41507223_0006

1. EINLEITUNG

1.1 Überblick über die Variable-Metrik-Methoden

Die meisten Verfahren zur Bestimmung eines lokalen Minimums einer zweifach stetig differenzierbaren Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ weisen die folgende Grundstruktur auf:

Ausgehend von einem Startvektor x_0 wird iterativ eine Folge von Punkten x_k erzeugt, welche sich dem gesuchten Minimum x^* nähern. In jedem Iterationsschritt wird zuerst eine Suchrichtung p_k bestimmt. Dann wird entlang dieser Richtung eine Schrittlänge λ_k ermittelt und zwar als geeignete Näherung für das Minimum von $f(x_k + \lambda p_k)$ ("line search").

Mit Variable-Metrik-Methode oder Quasi-Newton-Methode bezeichnet man eine wichtige Klasse von solchen Verfahren, deren wesentliches Merkmal darin besteht, dass zusätzlich eine Folge von Matrizen konstruiert wird, die zur Berechnung der Suchrichtung dienen.

Wählt man die Richtung mit der kleinsten Richtungsableitung zur Suchrichtung, so ist dies äquivalent zur Lösung der Aufgabe $\min \nabla f(x)^T p$ bezüglich $\|p\| = 1$. Daraus ergibt sich die Richtung $p = -\nabla f(x)$, sofern man in der Nebenbedingung die Euklidische Norm verwendet. Für die Norm $\|p\| := \sqrt{p^T M p}$ erhält man dagegen

$$p = \frac{-M^{-1} \nabla f(x)}{\sqrt{\nabla f(x)^T M^{-1} \nabla f(x)}}.$$

Wird nun die Matrix M in jeder Iteration geändert, so kann man zu Recht von einer Variable-Metrik-Methode sprechen.

In der Nähe eines Iterationspunktes x_k kann $\nabla f(x)$ aufgrund der Taylordarstellung durch $\nabla f(x_k) + \nabla^2 f(x_k)^T (x - x_k)$ angenähert werden. Da der Gradient im lokalen Minimum verschwindet, bestimmt man beim Newton-Verfahren den nächsten Iterationspunkt x_{k+1} als Lösung von

$$\nabla f(x_k) + \nabla^2 f(x_k)^T (x - x_k) = \vec{0}.$$

Sofern die Hesse-Matrix $\nabla^2 f(x_k)$ nichtsingulär ist, erhält man

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k).$$

In vielen Fällen ist die Berechnung und die Inversion der Hesse-Matrix zu aufwendig, und

es wird deshalb aus den Differenzen des Gradienten für $\nabla^2 f(x_k)^{-1}$ eine Approximation H_k konstruiert. In einem Verfahren mit obiger Grundstruktur werden dann Suchrichtungen der Form $p_k = -H_k \nabla f(x_k)$ verwendet. Durch diesen Ansatz wird der Namen Quasi-Newton-Methode begründet.

Von den Matrizen H_k verlangt man, dass sie symmetrisch sind und dass sie die folgende Gleichung erfüllen:

$$s_k = H_{k+1} y_k, \text{ wobei } s_k := x_{k+1} - x_k = \lambda_k p_k \text{ und } y_k := \nabla f(x_{k+1}) - \nabla f(x_k)$$

Diese zweite Bedingung wird *Quasi-Newton-Gleichung* genannt, im eindimensionalen Fall wird dadurch H_k^{-1} eindeutig als Sekantensteigung von $f'(x)$ festgelegt. In der Literatur wird eine grosse Anzahl von Formeln angegeben, welche unter Verwendung von H_k , s_k und y_k eine neue Matrix H_{k+1} generieren (vgl. Dennis und Moré [77]). Es ist üblich den Index k zu unterdrücken und $\bar{H}$ anstelle von H_{k+1} zu schreiben. Am bekanntesten sind die DFP-Formel (vgl. Davidon [59]), Fletcher und Powell [63])

$$\bar{H} = H + \frac{s s^T}{y^T s} - \frac{H y y^T H}{y^T H y}$$

und die BFGS-Formel (vgl. Broyden [69], Fletcher [70], Goldfarb [70], Shanno [70])

$$\begin{aligned} \bar{H} &= H + \left(1 + \frac{y^T H y}{y^T s} \right) \frac{s s^T}{y^T s} - \frac{s y^T H + H y s^T}{y^T s} \\ &= \left(I - \frac{s y^T}{y^T s} \right) H \left(I - \frac{y s^T}{y^T s} \right) + \frac{s s^T}{y^T s} . \end{aligned}$$

Beide Formeln erzeugen positiv definite Matrizen, sofern H_0 positiv definit ist und in jeder Iteration $y^T s > 0$ gilt.

Die folgende Beziehung stammt von Fletcher [70]:

$$\bar{H}_{BFGS} = \bar{H}_{DFP} + v v^T, \text{ wobei } v = \sqrt{y^T H y} \left(\frac{s}{y^T s} - H \frac{y}{y^T H y} \right)$$

Darüberhinaus zeigte er, dass eine ganze Klasse von Formeln, welche zuvor von Broyden angegeben wurde, durch $\bar{H}_\phi = \bar{H}_{DFP} + \phi v v^T$, $\phi \in \mathbb{R}$ dargestellt werden kann. Dixon [72] konnte für $\phi \geq 0$ nachweisen, dass die Folge x_k bei exaktem Line search unabhängig ist von der Wahl von ϕ . In der Praxis wird aber mit ungenauem Line search gearbeitet und

deshalb ergeben sich bei verschiedenen Parameterwerten ϕ numerisch völlig verschiedene Verfahren. Davidon [75] verwendet die freie Wahl von ϕ um die Konditionszahl von $H^{-1}\bar{H}_\phi$ minimal zu gestalten. Viele Autoren vertreten aber die Meinung, dass die BFGS-Formel die besten praktischen Eigenschaften besitzt (vgl. Brodlie [77], Shanno und Phua [78b]). Durch Greenstadt [70] wurde gezeigt, dass gewisse Formeln eine minimale Änderung der Matrix H bewirken. Der Unterschied der Matrizen $\bar{H}$ und H wird gemessen durch eine gewichtete Matrixnorm $\|M^{\frac{1}{2}}(\bar{H}-H)M^{\frac{1}{2}}\|_F$, wobei $\|A\|_F$ die Frobeniusnorm, i.e. $(\text{tr } A^T A)^{\frac{1}{2}}$, und M eine symmetrisch positiv definite Gewichtsmatrix bezeichnet. Als Lösung des Problems

$$\min \|M^{\frac{1}{2}}(\hat{H}-H)M^{\frac{1}{2}}\|_F \quad \text{bezüglich} \quad \hat{H} \in \mathbb{R}^n \times n, \hat{H} \text{ symmetrisch, } \hat{H}y = s$$

bestimmte Greenstadt eine Formel für $\bar{H}$. In der konkreten Realisierung wählte er für M die Einheitsmatrix beziehungsweise H^{-1} , doch haben die beiden resultierenden Formeln keine Verbreitung gefunden. Goldfarb [70] benützte die Lösung von Greenstadt setzte aber $M = \bar{H}^{-1}$ und erhielt dadurch die BFGS-Formel. Die Methode von Greenstadt ist von Dennis und Moré [77] neu formuliert und durch Dennis und Schnabel [79] erweitert worden.

1.2 Quadratische Näherungsfunktion

Das zentrale Problem bei den meisten Verfahren mit der oben angegebenen Grundstruktur ist das Bestimmen der Suchrichtung. Näherungsfunktionen sind ein mögliches Konzept dazu. Wir betrachten nun einen beliebigen aber festen Iterationspunkt x_k . Ausgehend von dieser Stelle ist ein Schritt in einer noch unbekannten Richtung $p \in \mathbb{R}^n$ auszuführen. Falls $f(x_k + p)$ in geeigneter Weise durch eine Funktion in p angenähert werden kann, so sprechen wir von einer Näherungsfunktion. In der englischen Fachliteratur wird dafür häufig der Begriff "Model" verwendet. Im folgenden schreiben wir vereinfachend x, f und g anstelle von $x_k, f(x_k)$ und $\nabla f(x_k)$. Die einfachste Näherungsfunktion ist die quadratische Funktion $q(p) = \beta + b^T p + \frac{1}{2} p^T B p$ mit $\beta \in \mathbb{R}$, $b \in \mathbb{R}^n$ und $B \in \mathbb{R}^n \times n$. Aus der Forderung

$$q(0) = f \quad \text{und} \quad \nabla q(0) = g \tag{1.2.1}$$

folgt sogleich $q(p) = f + g^T p + \frac{1}{2} p^T B p$. Die neue Suchrichtung $\tilde{p}$ soll nun als Minimum von $q(p)$ bestimmt werden, woraus sich $\tilde{p} = -B^{-1}g$ ergibt. Der Line search entlang dieser Richtung liefert den nächsten Iterationspunkt x_{k+1} . Auch hier schreiben wir vereinfachend $\bar{x}, \bar{f}, \bar{g}$ statt $x_{k+1}, f(x_{k+1}), \nabla f(x_{k+1})$.

Eine neue Näherungsfunktion $\bar{q}(p) = \bar{f} + \bar{g}^T p + \frac{1}{2} p^T \bar{B} p$ sollte nun an dieser Stelle bestimmt werden. Wie oben gilt $\bar{q}(0) = \bar{f}$ und $\nabla \bar{q}(0) = \bar{g}$. Zusätzlich kann man nun verlangen, dass die neue Näherungsfunktion auch am Startpunkt des eben ausgeführten Schrittes sowohl im Funktionswert wie auch im Gradient mit f übereinstimmt, das bedeutet:

$$\bar{q}(-s) = f, \quad (1.2.2)$$

$$\nabla \bar{q}(-s) = g, \quad (1.2.3)$$

wobei s wie oben für $\bar{x} - x$ steht. Aus (1.2.3) folgt unmittelbar

$$\bar{B}s = \bar{g} - g =: y. \quad (1.2.4)$$

Setzt man $\bar{H} = \bar{B}^{-1}$, so erkennt man, dass es sich bei (1.2.4) um die Quasi-Newton-Gleichung handelt. Durch Einsetzen von (1.2.4) in (1.2.2) erhält man die Bedingung

$$\bar{g}^T s + 2(f - \bar{f}) + g^T s = 0 \quad \text{oder} \quad \bar{f} - f - g^T s = f - \bar{f} + \bar{g}^T s \quad (1.2.5)$$

Es ist leicht einzusehen, dass eine quadratische Funktion f diese Gleichungen erfüllt, andere Funktionen im allgemeinen diese Bedingung aber verletzen. Dies bedeutet, dass bei den normalen Variable-Metrik-Verfahren zur Bestimmung der Suchrichtungen eigentlich nur die Gradientenwerte verwendet werden oder anders gesagt, dass nicht eine quadratische Näherungsfunktion für f sondern eine affine Näherungsfunktion für ∇f benutzt wird.

2. VERFAHREN MIT NICHTQUADRATISCHER NÄHERUNGSFUNKTION

2.1 Konische Näherungsfunktion

Mit der Gleichung (1.2.5) begründet Davidon [80] die Einführung von zwei Funktionen, es sind dies die kollineare Skalierung

$$x(w) = x + \frac{Jw}{h^T w + 1}$$

und die konische Funktion

$$c(p) = \beta + \frac{b^T p}{1 - a^T p} + \frac{\frac{1}{2} p^T B p}{(1 - a^T p)^2}$$

Davidon beschreibt neben den Eigenschaften dieser Funktionen zwei Algorithmen, welche zur Bestimmung der Suchrichtung Funktions- und Gradientenwerte von n aufeinanderfolgenden Iterationspunkten verwenden.

Sorensen [80] verwendet in einer Variable-Metrik-Methode eine quadratische Näherungsfunktion unterwirft aber vorher den Vektor x einer kollinearen Skalierung

$$x(w) = x + \frac{Jw}{h^T w + 1}$$

mit $w, h \in \mathbb{R}^n$ und $J \in \mathbb{R}^{n \times n}$ regulär. Genauer gesagt bedeutet dies, dass $f(x(w))$ durch eine quadratische Näherungsfunktion $q(w) = \beta + b^T w + \frac{1}{2} w^T B w$ approximiert wird. Fordert man analog zu (1.2.1) die Bedingungen

$$q(0) = f(x(0)) \quad \text{und} \quad \nabla_w q(0) = \nabla_x f(x(0)) ,$$

so folgt aus der Kettenregel $\nabla_w f(x(w)) = \nabla_w x(w) \nabla_x f(x(w))$ und wegen

$$\nabla_w x(w) = \frac{1}{h^T w + 1} J^T \left(I - \frac{h w^T}{h^T w + 1} \right)$$

die Darstellung

$$q(w) = f + g^T J w + \frac{1}{2} w^T B w . \quad (2.1.1)$$

Bei diesem Ansatz werden, gemäss der Definition von $z(w)$, Suchrichtungen mit der Gestalt

$$p = \frac{J w}{h^T w + 1}$$

benützt. Diese Gleichung kann nach w aufgelöst werden, indem die Inverse von $J - p h^T$, sofern sie existiert, nach der Sherman-Morrison-Formel (vgl. z.B. Ortega und Rheinboldt [70] p 50) berechnet wird:

$$w = \frac{J^{-1} p}{1 - h^T J^{-1} p}$$

Wird nun dieser Ausdruck in (2.1.1) eingesetzt und zugleich die Matrix $(J^{-1})^T B J^{-1}$ durch A und der Vektor $(J^{-1})^T h$ durch a abgekürzt, so erhält man für f die folgende konische Näherungsfunktion

$$c(p) = f + \frac{g^T p}{1 - a^T p} + \frac{\frac{1}{2} p^T A p}{(1 - a^T p)^2} \quad (2.1.2)$$

Das Verfahren, welches Sorensen beschrieben hat, kann deshalb auch als Variable-Metrik-Methode mit konischer Näherungsfunktion aufgefasst werden. Im folgenden wird das Verfahren ausführlich in dieser zweiten Form dargestellt, denn nur so kann seine Geometrie klar erfasst werden.

Die konische Funktion aus (2.1.2) kann in die folgende Gestalt gebracht werden

$$c(p) = \frac{1}{2} \left(g + \frac{A p}{1 - a^T p} \right)^T A^{-1} \left(g + \frac{A p}{1 - a^T p} \right) + f - \frac{1}{2} g^T A^{-1} g ,$$

dadurch lässt sich der Gradient leicht berechnen

$$\nabla c(p) = \frac{1}{1 - a^T p} \left(I + \frac{a p^T}{1 - a^T p} \right) \left(g + \frac{A p}{1 - a^T p} \right) . \quad (2.1.3)$$

Im Iterationspunkt z wird eine neue Suchrichtung $\tilde{p}$ bestimmt und zwar wie bei der quadratischen Näherungsfunktion als Minimum von $c(p)$. Folglich muss $\nabla c(\tilde{p}) = 0$ gelten. Dies ist nur möglich, wenn in (2.1.3) die hintere Klammer verschwindet. Die dadurch entstehende Gleichung löst man nach $\tilde{p}$ auf und erhält

$$\tilde{p} = \frac{-A^{-1} g}{1 - a^T A^{-1} g} . \quad (2.1.4)$$

Damit $\tilde{p}$ auch tatsächlich ein Minimum ist, müssen wir noch die zweiten Ableitungen von $c(p)$ untersuchen

$$\nabla^2 c(p) = \frac{\nabla c(p) a^T + a \nabla c(p)^T}{1 - a^T p} + \frac{1}{(1 - a^T p)^2} \left(I + \frac{a p^T}{1 - a^T p} \right)^T A \left(I + \frac{a p^T}{1 - a^T p} \right).$$

An der kritischen Stelle $\tilde{p}$ wird dieser Ausdruck zu

$$\nabla^2 c(\tilde{p}) = \left(1 - a^T A^{-1} g \right)^2 \left(I - a g^T A^{-1} \right)^T A \left(I - a g^T A^{-1} \right).$$

Diese Matrix ist positiv definit und damit $\tilde{p}$ ein Minimum von $c(p)$, sofern A positiv definit ist und $1 - a^T A^{-1} g \neq 0$ gilt. Bei den gewöhnlichen Variable-Metrik-Methoden werden die Matrizen positiv definit gehalten um ausschliesslich Abstiegsrichtungen (i.e. $g^T \tilde{p} < 0$) zu erzeugen. Auch bei Verwendung von konischen Näherungsfunktionen gelingt es nur positiv definite Matrizen A zu generieren, die Abstiegeigenschaft ist jedoch zusätzlich vom Vorzeichen von $1 - a^T A^{-1} g$ abhängig. Wir werden später auf dieses Problem zurückkommen.

Durch einen geeigneten Line search entlang der Richtung $\tilde{p}$ wird der neue Iterationspunkt $\bar{x}$ bestimmt, womit auch der Funktionswert $\bar{f}$ und der Gradient $\bar{g}$ berechnet werden kann. Wählt man entsprechend zu (2.1.2) als neue Näherungsfunktion

$$\bar{c}(p) = \bar{f} + \frac{\bar{g}^T p}{1 - \bar{a}^T p} + \frac{\frac{1}{2} p^T \bar{A} p}{(1 - \bar{a}^T p)^2} \quad (2.1.5)$$

so gilt offensichtlich $\bar{c}(0) = \bar{f}$ und $\nabla \bar{c}(0) = \bar{g}$. Analog zu (1.2.2) und (1.2.3) verlangen wir, dass $\bar{c}(p)$ die beiden folgenden Bedingungen erfüllt

$$\bar{c}(-s) = f \quad (2.1.6)$$

$$\nabla \bar{c}(-s) = g. \quad (2.1.7)$$

Wenn man zur Vereinfachung die Grösse

$$\gamma := \frac{1}{1 + \bar{a}^T s} \quad (2.1.8)$$

eingführt, so folgt aus (2.1.7) unter Verwendung von $(I - \gamma \bar{a} s^T)^{-1} = I + \bar{a} s^T$ eine spezielle Form der Quasi-Newton-Gleichung

$$\gamma^2 \bar{A} s = \gamma \bar{g} - g - g^T s \bar{a} =: y. \quad (2.1.9)$$

Durch Einsetzen in (2.1.6) erhält man

$$\gamma \bar{g}^{Ts} + 2(f - \bar{f}) + \frac{1}{\gamma} g^{Ts} = 0 . \quad (2.1.10)$$

Diese Gleichung ist nahe verwandt mit (1.2.5) und liefert

$$\gamma = \frac{f - \bar{f} \pm \rho}{-\bar{g}^{Ts}} = \frac{-g^{Ts}}{f - \bar{f} \mp \rho} ,$$

wobei wir die folgende Abkürzung benützen

$$\rho := \sqrt{(f - \bar{f})^2 - g^{Ts} \bar{g}^{Ts}} . \quad (2.1.11)$$

Damit γ und ρ überhaupt existieren, darf der Ausdruck unter der Wurzel nicht negativ sein. Durch den Line search lässt sich dies leicht erreichen, falls ausschliesslich Abstiegsrichtungen verwendet werden. Um zu entscheiden, welches der Zeichen $\pm$ man bei γ verwenden soll, betrachten wir die Hyperebene festgelegt durch $1 - \bar{a}^T p = 0$, über der die Funktion $\bar{c}(p)$ nicht definiert ist. Damit die Punkte x und $\bar{x}$ bezüglich dieser Ebene im selben Halbraum liegen, muss $\gamma > 0$ erfüllt sein, woraus sich

$$\gamma = \frac{f - \bar{f} - \rho}{-\bar{g}^{Ts}} = \frac{-g^{Ts}}{f - \bar{f} + \rho} \quad (2.1.12)$$

ergibt. Wie man leicht einsieht, gilt wegen der Bemerkung zu (1.2.5) für quadratische Funktionen f stets $\gamma = 1$. Im Iterationspunkt $\bar{x}$ sind alle Informationen vorhanden um γ zu berechnen, aus (2.1.8) folgt deshalb für $\bar{a}$ die Bedingung

$$\bar{a}^{Ts} = \frac{1}{\gamma} - 1 . \quad (2.1.13)$$

Die Variable-Metrik-Methode mit konischer Näherungsfunktion erhält somit die folgende Gestalt:

- | | |
|------------------|---|
| <i>Schritt 1</i> | Wähle x_0 und A_0 .
Berechne f_0 , g_0 und $p_0 = -A_0^{-1} g_0$.
Setze $k = 0$. |
| <i>Schritt 2</i> | Bestimme durch einen Line search $x_{k+1} = x_k + \lambda_k p_k$,
sodass $(f_k - f_{k+1})^2 - g_k^{Ts_k} g_{k+1}^{Ts_k} > 0$ gilt.
Teste die Abbruchbedingungen. |
| <i>Schritt 3</i> | Berechne γ_k nach der Formel (2.1.12).
Wähle einen Vektor a_{k+1} , welcher (2.1.13) erfüllt, |

und berechne y_k nach der Formel (2.1.9).

Bestimme nach einer der bekannten Formeln (vgl. Kapitel 1.1) eine Matrix A_{k+1} , welche (2.1.9) erfüllt.

Schritt 4 Berechne p_{k+1} nach der Formel (2.1.4) und setze $k = k + 1$.
Gehe zurück zu Schritt 2.

Wie in Kapitel 1.1 dargestellt wurde, generieren die wichtigsten Update-Formeln ausschliesslich positiv definite Matrizen, sofern $y^T s > 0$ gilt. Für den Vektor y definiert in (2.1.9) folgt aus (2.1.8) und (2.1.12)

$$y^T s = \gamma \bar{g}^T s - \frac{1}{\gamma} g^T s = 2\rho > 0, \quad (2.1.14)$$

womit wir durch die entsprechenden Formeln auch bei konischen Näherungsfunktionen positiv definite Matrizen erhalten.

Beim eben dargestellten Algorithmus bleibt vorläufig unklar, wie der Vektor $\bar{a}$ in Schritt 3 bestimmt werden soll. Wählt man beispielsweise

$$\bar{a} = \frac{1}{g^T s} \left(\gamma \bar{g} - g - \frac{2\rho}{\|s\|^2} s \right)$$

so ist die Bedingung (2.1.13) erfüllt, und aus (2.1.9) erhalten wir $y = \frac{2\rho}{\|s\|^2} s$. Die Quasi-Newton-Gleichung reduziert sich auf $\gamma^2 \bar{A} s = \frac{2\rho}{\|s\|^2} s$. Zur Erzeugung der Matrizen $\bar{A}$ könnte man daher die Formel

$$\bar{A} = \frac{2\rho}{\|s\|^2} I$$

verwenden, wodurch ein Verfahren des steilsten Abstiegs resultiert. Die guten Eigenschaften der Variable-Metrik-Methoden bleiben so nicht erhalten.

Um einen derart ungünstigen Vektor $\bar{a}$ zu verhindern, versuchen wir den Einfluss von $\bar{a}$ möglichst gering zu halten. Das heisst wir bestimmen $\bar{a}$ als Lösung von

$$\min \| \bar{a} \|^2 \quad \text{bezüglich} \quad \bar{a}^T s = \frac{1}{\gamma} - 1.$$

Dieses Problem ist abhängig von der verwendeten Norm, und entsprechend der variablen Metrik betrachten wir hier $\|x\| := \sqrt{x^T M x}$. Für die euklidische Norm, das heisst

$M = I$, folgt aus der Lagrange-Bedingung

$$\bar{a} = \left(\frac{1}{\gamma} - 1 \right) \frac{s}{s^T s} \quad \text{und} \quad y = \gamma \bar{g} - g - \left(\frac{1}{\gamma} - 1 \right) \frac{g^T s}{s^T s} s. \quad (2.1.15)$$

In Analogie zur Methode von Greenstadt, welche wir in Kapitel 1.1 dargestellt haben, wählen wir nun $M = A^{-1}$, woraus sich

$$\bar{a} = \left(\frac{1}{\gamma} - 1 \right) \frac{1}{g^T s} g \quad \text{und} \quad y = \gamma \bar{g} - \frac{1}{\gamma} g \quad (2.1.16)$$

ergibt. Sorensen [80] arbeitet mit diesem Vektor y ohne dafür aber eine Begründung anzugeben. In Anlehnung an Goldfarb setzen wir nun noch $M = \bar{A}^{-1}$ und erhalten

$$\bar{a} = \frac{\frac{1}{\gamma} - 1}{\gamma \bar{g}^T s - g^T s} \left(\gamma \bar{g} - g \right) \quad \text{und} \quad y = \frac{2\rho}{\gamma \bar{g}^T s - g^T s} \left(\gamma \bar{g} - g \right). \quad (2.1.17)$$

Eine andere ähnliche Strategie versucht die Abweichung von $\bar{g} - g$ zu y möglichst klein zu machen. In diesem Fall berechnet man den Vektor y als Lösung von

$$\min \| \bar{g} - g - y \|^2 \quad \text{bezüglich} \quad y^T s = 2\rho.$$

Wieder kann dabei die Norm auf verschiedene Arten gewählt werden. Für die euklidische Norm ergibt sich

$$y = \bar{g} - g + \frac{2\rho - \bar{g}^T s + g^T s}{s^T s} s. \quad (2.1.18)$$

Benützt man dagegen in der Norm die Matrix A^{-1} , so folgt

$$y = \bar{g} - \frac{\bar{g}^T s - 2\rho}{g^T s} g. \quad (2.1.19)$$

Bei der $\bar{A}^{-1}$ -Norm erhalten wir

$$y = \frac{2\rho}{\bar{g}^T s - g^T s} \left(\bar{g} - g \right). \quad (2.1.20)$$

Diese Möglichkeiten bei der Wahl von $\bar{a}$ beziehungsweise y können einheitlich dargestellt werden:

$$y = \gamma \bar{g} - \frac{1}{\gamma} g + \left(\frac{1}{\gamma} - 1 \right) \left(\mu (\bar{g}^T s \, g - g^T s \, \bar{g}) + \nu (s^T s \, g - g^T s \, s) \right). \quad (2.1.21)$$

Dabei bezeichnen μ und ν reelle Parameter. Die beiden Klammern hinter μ und ν stellen Vektoren orthogonal zu s dar, während der verbleibende Ausdruck uns gemäss (2.1.14) garantiert, dass $y^T s = 2\rho$ gilt. Für die betrachteten Formeln fassen wir die zugehörigen Parameter μ und ν in der folgenden Tabelle zusammen

Formel	Bezeichnung	μ	ν
(2.1.15)	con1	0	$\frac{1}{s^T s}$
(2.1.16)	con2	0	0
(2.1.17)	con3	$\frac{1}{\bar{g}^T s - \frac{1}{\gamma} g^T s}$	0
(2.1.18)	con4	$-\frac{\gamma}{g^T s}$	$\left(\frac{\gamma \bar{g}^T s}{g^T s} + 1 \right) \frac{1}{s^T s}$
(2.1.19)	con5	$-\frac{\gamma}{g^T s}$	0
(2.1.20)	con6	$\frac{1+\gamma}{\bar{g}^T s - g^T s}$	0

Für den Vektor $\bar{a}$ erhält man aus (2.1.9) und (2.1.21) die Darstellung

$$\bar{a} = \left(\frac{1}{\gamma} - 1 \right) \left(\frac{1}{g^T s} g - \mu \left(\frac{\bar{g}^T s}{g^T s} g - \bar{g} \right) - \nu \left(\frac{s^T s}{g^T s} g - s \right) \right). \quad (2.1.22)$$

Es bleibt nun noch eine offene Frage zu untersuchen: Die neue Suchrichtung definiert gemäss (2.1.4) ist keine Abstiegsrichtung, falls $1 - \bar{a}^T \bar{A}^{-1} \bar{g} < 0$ gilt. Um die geometrische Bedeutung dieser Ungleichung zu studieren, betrachten wir unsere konischen Funktionen $\bar{c}(p)$ entlang der Suchrichtung, das heisst für Vektoren $p = -\lambda \bar{A}^{-1} \bar{g}$. Zusammen mit der Definition

$$\psi(\lambda) := \frac{\frac{1}{\lambda} + \bar{a}^T \bar{A}^{-1} \bar{g} - \frac{1}{2}}{\left(\frac{1}{\lambda} + \bar{a}^T \bar{A}^{-1} \bar{g}\right)^2}$$

erhalten wir $\bar{c}(\lambda) = \bar{f} - \psi(\lambda) \cdot \bar{g}^T \bar{A}^{-1} \bar{g}$. Wie oben gezeigt wurde, beschreibt

$$\lambda_1 = \frac{1}{1 - \bar{a}^T \bar{A}^{-1} \bar{g}}$$

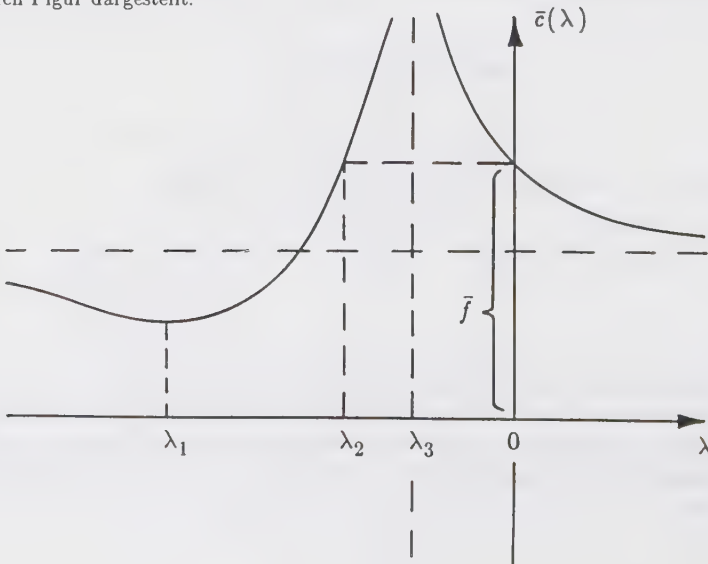
ein lokales Minimum von $\bar{c}(\lambda)$ mit dem Funktionswert $\bar{c}(\lambda_1) = \bar{f} - \frac{1}{2} \bar{g}^T \bar{A}^{-1} \bar{g}$. Es ist leicht einzusehen, dass $\psi(\lambda) \leq \frac{1}{2}$ gilt, und deshalb ist λ_1 sogar ein globales Minimum von $\bar{c}(\lambda)$. Für

$$\lambda_2 = \frac{1}{\frac{1}{2} - \bar{a}^T \bar{A}^{-1} \bar{g}}$$

folgt $\bar{c}(\lambda) = \bar{f}$, und für

$$\lambda_3 = \frac{1}{-\bar{a}^T \bar{A}^{-1} \bar{g}}$$

ist $\bar{c}(\lambda)$ nicht definiert, mit dieser Schrittweite trifft unsere Suchrichtung auf die Hyperebene $1 - \bar{a}^T p = 0$. Wir betrachten nun den kritischen Fall $\lambda_1 < 0$, woraus wir unmittelbar $\lambda_1 < \lambda_2 < \lambda_3 < 0$ erhalten. Die zugrunde liegende Situation ist in der folgenden Figur dargestellt.



Das Minimum bestimmt durch λ_1 liegt in einem anderen Halbraum bezüglich der Hyperebene $1 - \bar{a}^T p = 0$ als die Punkte x und $\bar{x}$. Für $\lambda > 0$ ist $\psi'(\lambda)$ positiv und somit $c(\lambda)$ stets abnehmend. Aus $\lambda \rightarrow \infty$ ergibt sich zudem

$$\bar{c}(\lambda) \rightarrow \bar{f} - \frac{\bar{a}^T \bar{A}^{-1} \bar{g} - \frac{1}{2}}{(\bar{a}^T \bar{A}^{-1} \bar{g})^2} \bar{g}^T \bar{A}^{-1} \bar{g}.$$

Offensichtlich darf man in diesem Fall die Suchrichtung nicht mit der anfänglichen Schrittlänge λ_1 versehen. Aufgrund der Näherungsfunktion sollte man stattdessen dafür eine hinreichend grosse, positive Zahl verwenden. Für die Variante *con3*, d. h. $\bar{a}$ wird gemäss (2.1.17) gewählt, errechnet man

$$1 - \bar{a}^T \bar{A}^{-1} \bar{g} = \frac{1}{2\rho} (y^T s - \bar{a}^T s y^T \bar{A}^{-1} \bar{g}) = \frac{1}{2\rho} (\gamma^2 \bar{g}^T s - \frac{1}{\gamma} g^T s).$$

Sobald $\gamma^3 \bar{g}^T s - g^T s > 0$ gilt, ist deshalb stets $\lambda_1 > 0$ erfüllt. Durch einen hinreichen genauen Line search kann diese Ungleichung leicht garantiert werden. Dieselbe Aussage kann auch gemacht werden, wenn $\bar{a}$ parallel zu $\gamma \bar{g} - \frac{1}{\gamma} g$ gewählt wird.

2.2 Rationale Näherungsfunktion

Barzilai und Ben-Tal [82] betrachten Algorithmen für die Minimierung von Funktionen einer Variablen (Line search) und verwenden dabei die folgende Näherungsfunktion

$$f(x) = \frac{ax^2 + bx + c}{dx - 1},$$

welche sie zur Unterscheidung mit dem Adjektiv *rational* versehen. Im folgenden werden Variable-Metrik-Methoden studiert, welche solche rationalen Näherungsfunktionen benutzen, das heisst Funktionen der Form

$$r(p) = f + \frac{g^T p + \frac{1}{2} p^T A p}{1 - a^T p}. \quad (2.2.1)$$

Durch Differenzieren der Gleichung $r(p)(1 - a^T p) = f(1 - a^T p) + g^T p + \frac{1}{2} p^T A p$ lässt sich der Gradient dieser Funktion leicht berechnen

$$\nabla r(p) = \frac{1}{1 - a^T p} \left(g + A p + (r(p) - f) a \right). \quad (2.2.2)$$

Wie bei den anderen Näherungsfunktionen bestimmen wir im Iterationspunkt x die neue Suchrichtung $\tilde{p}$ als Minimum von $r(p)$. Aus der Bedingung $\nabla r(\tilde{p}) = 0$ folgt $\tilde{p} = -A^{-1}g - (r(\tilde{p}) - f)A^{-1}a$, doch ist der gesuchte Vektor leider noch immer in der rechten Seite dieser Gleichung enthalten. Durch Einführen der Grösse $\mu := f - r(\tilde{p})$ ergibt sich

$$\tilde{p} = -A^{-1}g + \mu A^{-1}a. \quad (2.2.3)$$

Indem wir diesen Ausdruck in die Definition von μ einsetzen, erhalten wir die folgende quadratische Gleichung $a^T A^{-1}a \mu^2 - 2(1 + a^T A^{-1}g)\mu + g^T A^{-1}g = 0$ mit den Lösungen

$$\mu = \frac{1 + a^T A^{-1}g}{a^T A^{-1}a} \pm \sqrt{\left(\frac{1 + a^T A^{-1}g}{a^T A^{-1}a}\right)^2 - \frac{g^T A^{-1}g}{a^T A^{-1}a}}.$$

Die Matrix der zweiten Ableitung $\nabla^2 r(p)$ hat die Gestalt

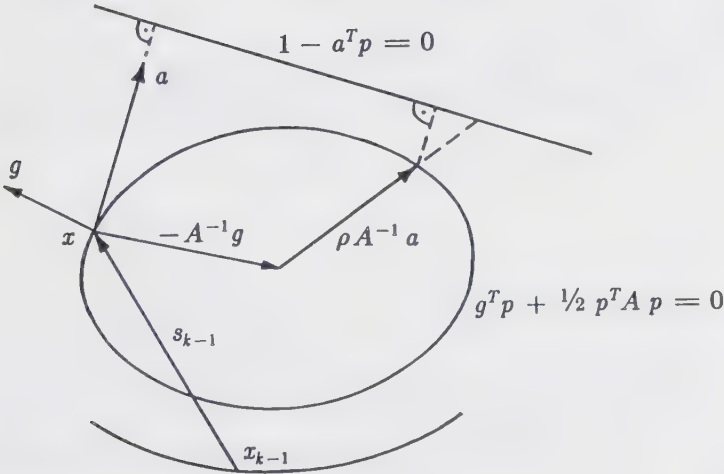
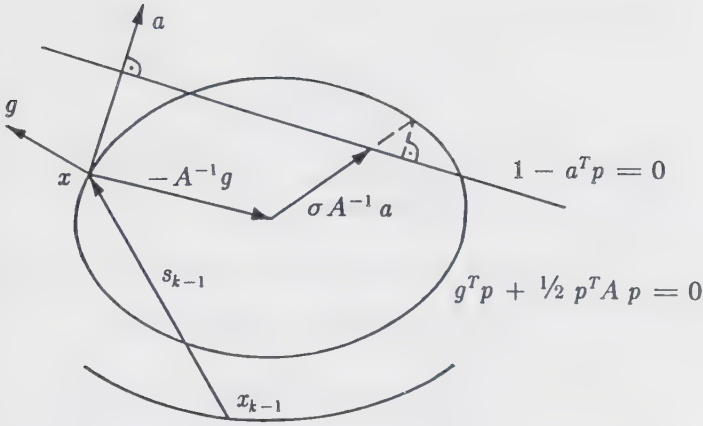
$$\frac{1}{1 - a^T p} \left(A + a \nabla r(p)^T + \nabla r(p) a^T \right)$$

und reduziert sich an der Stelle $\tilde{p}$ zu $(1 - a^T \tilde{p})^{-1} A$. Damit tatsächlich ein Minimum vorliegt, muss die Matrix A positiv definit gehalten werden und $1 - a^T \tilde{p} > 0$ gelten. Diese Ungleichung ist wegen (2.2.3) und dem obigen Ausdruck für μ äquivalent zu $\mp \sqrt{(1 + a^T A^{-1}g)^2 - g^T A^{-1}g a^T A^{-1}a} > 0$, woraus wir

$$\mu = \frac{1 + a^T A^{-1}g}{a^T A^{-1}a} - \sqrt{\left(\frac{1 + a^T A^{-1}g}{a^T A^{-1}a}\right)^2 - \frac{g^T A^{-1}g}{a^T A^{-1}a}}. \quad (2.2.4)$$

erhalten.

Offensichtlich hat unsere Näherungsfunktion $r(p)$ kein Minimum, wenn der Term unter der Wurzel negativ ist. Um diese Situation aufzuklären, werden nun Zählerfunktion und Nennerfunktion einzeln untersucht, besonders interessieren ihre Hyperflächen vom Niveau Null. Wir unterscheiden zwei Fälle, je nachdem ob die Nullhyperebene von $1 - a^T p$ das Nullellipsoid von $g^T p + \frac{1}{2} p^T A p$ schneidet oder nicht. Bewegt man sich im *ersten Fall* vom Mittelpunkt des Ellipsoids derart nach aussen, dass man zuerst auf die Nullhyperebene trifft, so strebt die Funktion r entlang dieses Weges gegen $-\infty$. Bewegt man sich aber so vom Mittelpunkt nach aussen, dass man zuerst auf das Nullellipsoid trifft, dann strebt r gegen $+\infty$. Wie man sich im *zweiten Fall* vom Ellipsoidmittelpunkt auch wegbewegt, immer strebt die Funktion r gegen $+\infty$.



Um die beiden Fälle nun auch quantitativ zu unterscheiden, betrachten wir die Suchrichtung definiert durch (2.2.3) und denken uns μ als Variable. Wie man leicht nachrechnet, wird durch den Variablenwert

$$\sigma := \frac{1 + a^T A^{-1} g}{a^T A^{-1} a} \quad (2.2.5)$$

der Durchstosspunkt mit der Nullhyperebene und durch

$$\rho := \sqrt{\frac{g^T A^{-1} g}{a^T A^{-1} a}} \quad (2.2.6)$$

der Durchstosspunkt mit dem Nullellipsoid festgelegt. Für unser Verfahren ist der zweite Fall aus der obigen Unterscheidung günstig, er lässt sich durch

$$\sigma > \rho \quad (2.2.7)$$

umschreiben. Auch die Formel (2.2.4) kann einfacher dargestellt werden:

$$\mu = \sigma - \sqrt{\sigma^2 - \rho^2} \quad (2.2.8)$$

Über die Abstiegs Eigenschaft der Suchrichtung $\tilde{p}$ gibt das folgende Lemma Auskunft.

Lemma 2.2.1: Falls (2.2.7) erfüllt ist, so wird durch (2.2.3) und (2.2.4) eine Abstiegsrichtung definiert.

Beweis: Die Ungleichung (2.2.7) impliziert $\mu > 0$. Durch Multiplikation mit

$$-\sqrt{\sigma^2 - \rho^2} - \frac{1}{a^T A^{-1} a}$$

erhält man daraus

$$-\sigma \sqrt{\sigma^2 - \rho^2} + \sigma^2 - \rho^2 - \frac{\mu}{a^T A^{-1} a} = -\rho^2 + \mu \left(\sigma - \frac{1}{a^T A^{-1} a} \right) < 0.$$

Multipliziert man nun diese Ungleichung mit $a^T A^{-1} a$, so ergibt sich $g^T \tilde{p} < 0$. □

Auf das Problem, wie nun der Vektor a zu wählen ist, werden wir am Ende des Kapitels zurückkommen.

Sobald der nächste Iterationspunkt $\bar{x}$ durch den Line search bestimmt ist und die neuen Werte $\bar{f}$ und $\bar{g}$ berechnet sind, soll eine neue Näherungsfunktion

$$\bar{r}(p) = \bar{f} + \frac{\bar{g}^T p + \frac{1}{2} p^T \bar{A} p}{1 - \bar{a}^T p} \quad (2.2.9)$$

erzeugt werden. Entsprechend den Bedingungen (1.2.2) und (1.2.3) fordern wir

$$\bar{r}(-s) = f \quad (2.2.10)$$

$$\nabla \bar{r}(-s) = g \quad (2.2.11)$$

Zur Vereinfachung definieren wir wie bei der konischen Näherungsfunktion

$$\gamma := \frac{1}{1 + \bar{a}^T s} \quad (2.2.12)$$

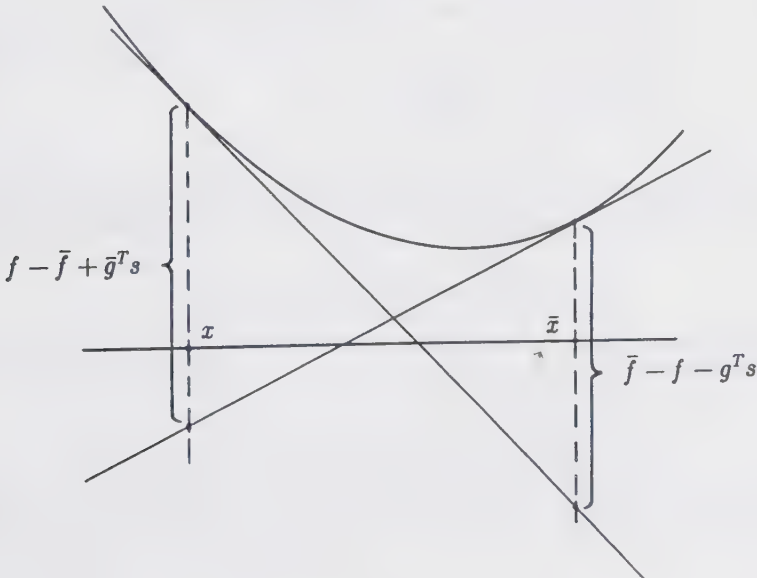
und erhalten aus (2.2.11) die folgende Quasi-Newton-Gleichung

$$\bar{A} s = \bar{g} - \frac{1}{\gamma} g + (f - \bar{f}) \bar{a} =: y \quad (2.2.13)$$

Wird nun dieser Ausdruck in (2.2.10) eingesetzt, so folgt $(f - \bar{f} + \bar{g}^T s) \gamma = \bar{f} - f - g^T s$
- man vergleiche wieder mit (1.2.5) - und deshalb gilt

$$\gamma = \frac{\bar{f} - f - g^T s}{f - \bar{f} + \bar{g}^T s} \quad (2.2.14)$$

Es ist auch hier zu bemerken, dass wir für quadratische Funktionen f stets $\gamma = 1$ erhalten. Damit die Punkte x und $\bar{x}$ bezüglich der Ebene $1 - \bar{a}^T p = 0$ im selben Halbraum liegen, sollte $\gamma > 0$ erfüllt sein.



Die obenstehende Figur zeigt, wie Zähler und Nenner von (2.2.14) geometrisch interpretiert werden können. Für konvexe Funktionen gilt immer $\gamma > 0$, sobald jedoch f in der Suchrichtung eine negative Krümmung aufweist, kann der Zähler von (2.2.14) negativ werden. In dieser Situation kann die Schrittlänge verlängert werden bis

$$\bar{f} - f - g^T s > 0 \quad \text{und} \quad f - \bar{f} + \bar{g}^T s > 0 \quad (2.2.15)$$

gilt. Dies funktioniert natürlich nur, falls f nach unten beschränkt ist, durch Ersetzen von $f(x)$ durch $e^{f(x)}$ kann solches aber immer erreicht werden.

Nun kann der Algorithmus kurz zusammengefasst werden:

- Schritt 1* Wähle x_0 und A_0 .
 Berechne f_0 , g_0 und $p_0 = -A_0^{-1} g_0$.
 Setze $k = 0$.
- Schritt 2* Bestimme durch einen Line search $x_{k+1} = x_k + \lambda_k p_k$, sodass (2.2.15) erfüllt ist. Teste die Abbruchbedingungen.
- Schritt 3* Berechne γ_k nach der Formel (2.2.14).
 Wähle einen Vektor a_{k+1} mit $a_{k+1}^T s_k = \frac{1}{\gamma_k} - 1$
 und berechne y_k nach der Formel (2.2.13).
 Bestimme nach einer der bekannten Formeln (vgl. Kapitel 1.1) eine Matrix A_{k+1} , welche (2.2.13) erfüllt.
- Schritt 4* Berechne μ_{k+1} und p_{k+1} nach den Formeln (2.2.4) und (2.2.3) und setze $k = k + 1$.
 Gehe zurück zu Schritt 2.

Damit die wichtigsten Formeln nur positiv definite Matrizen A erzeugen, müssen wir noch die Ungleichung $y^T s > 0$ überprüfen. Aus (2.2.13), der Definition von y , erhält man

$$y^T s = \frac{1}{\gamma} (f - \bar{f}) + \bar{g}^T s + (\bar{f} - f) - \frac{1}{\gamma} g^T s$$

und durch Umformen von (2.2.14) ergibt sich $\frac{1}{\gamma} (f - \bar{f}) + \bar{g}^T s = (\bar{f} - f) - \frac{1}{\gamma} g^T s$, zusammen folgt deshalb

$$\frac{1}{2} y^T s = \frac{1}{\gamma} (f - \bar{f}) + \bar{g}^T s = (\bar{f} - f) - \frac{1}{\gamma} g^T s = \frac{(f - \bar{f})^2 - \bar{g}^T s g^T s}{\bar{f} - f - g^T s}. \quad (2.2.16)$$

Indem man die Schrittlänge verlängert, kann im Line search $y^T s > 0$ garantiert werden.

Im folgenden wird dargestellt wie der Vektor $\bar{a}$ konkret realisiert werden kann.

Lemma 2.2.2 : Wenn $\bar{a}$ parallel zu $\bar{g} - g$ gewählt wird, dann ist $\bar{\sigma} > \bar{\rho}$ erfüllt.

Beweis : Es sei $\bar{a}$ so gewählt, dass

$$\bar{g} = \bar{\sigma} \bar{a} - \frac{1}{\bar{a}^T s} y \quad (2.2.17)$$

gilt. Durch Multiplikation mit $\bar{a}^T \bar{A}^{-1}$ erhält man aus (2.2.17) tatsächlich die Gleichung (2.2.5). Durch Umformen von (2.2.17) ergibt sich dagegen $(\frac{1}{\gamma} - 1) \bar{g} + \bar{g} - \frac{1}{\gamma} g + (f - \bar{f}) \bar{a} = \bar{a}^T s \bar{\sigma} \bar{a}$, und aus dieser Gleichung folgt, dass $\bar{g} - g$ identisch ist mit $\gamma(\bar{a}^T s \bar{\sigma} + \bar{f} - f) \bar{a}$. Multipliziert man nun (2.2.17) noch mit s , so kann man $\bar{\sigma}$ explizit berechnen

$$\bar{\sigma} = \frac{y^T s}{\bar{a}^T s^2} + \frac{\bar{g}^T s}{\bar{a}^T s}.$$

Die Gleichung

$$\bar{g}^T \bar{A}^{-1} \bar{g} = \bar{\sigma}^2 \bar{a}^T \bar{A}^{-1} \bar{a} - 2\bar{\sigma} \frac{\bar{a}^T \bar{A}^{-1} y}{\bar{a}^T s} + \frac{y^T \bar{A}^{-1} y}{\bar{a}^T s^2} \quad (2.2.18)$$

erhält man ebenfalls aus (2.2.17). Die beiden hinteren Summanden von (2.2.18) können zusammengefasst werden zu

$$-\left(2\bar{\sigma} - \frac{y^T s}{\bar{a}^T s^2}\right).$$

Wir zeigen nun, dass die obige Klammer positiv ist, woraus $\bar{\sigma} > 0$ folgt, zusammen mit (2.2.18) ergibt sich dann $\bar{\sigma} > \bar{\rho}$. Der Ausdruck

$$2\bar{\sigma} - \frac{y^T s}{\bar{a}^T s^2} = \frac{y^T s}{\bar{a}^T s^2} + \frac{2\bar{g}^T s}{\bar{a}^T s}$$

entspricht wegen (2.2.16)

$$\frac{2}{\bar{a}^T s^2} \left(\frac{1}{\gamma} (f - \bar{f}) + \bar{g}^T s + (\frac{1}{\gamma} - 1) \bar{g}^T s \right).$$

Der Term in der Klammer reduziert sich zu $\frac{1}{\gamma} (f - \bar{f} + \bar{g}^T s)$, was gemäss (2.2.15) positiv ist. $\square$

3. KONVERGENZEIGENSCHAFTEN

3.1 Lokale Konvergenz bei Variable-Metrik-Verfahren

Von Broyden, Dennis und Moré [73] stammen Untersuchungen zur lokalen Konvergenz von Variable-Metrik-Methoden mit quadratischer Näherungsfunktion, welche hier auszugsweise wiedergegeben werden, um aufzuzeigen wo der Beweisgang für konische bzw. rationale Näherungsfunktionen abgeändert werden muss. Zuerst formulieren wir einige Voraussetzungen, welche wir im folgenden benötigen:

1. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ zweifach stetig differenzierbar in einer offenen konvexen Menge D .
2. Sei $\nabla^2 f(x)$ Lipschitz-stetig in D ,
dh. $\|\nabla^2 f(x) - \nabla^2 f(\bar{x})\| \leq \kappa \|x - \bar{x}\|$ für $x, \bar{x} \in D$ und eine reelle Konstante κ .
3. Sei $x^* \in D$ ein lokales Minimum von f mit $\nabla^2 f(x^*)$ nichtsingulär,
dh. D kann eingeschränkt werden, sodass $\pi \|z\|^2 \leq z^T \nabla^2 f(x) z \leq \Pi \|z\|^2$ gilt für zwei positive reelle Konstanten Π und π sowie für $x \in D$ und $z \in \mathbb{R}^n$.
O.B.d.A darf man $\Pi = \kappa$ voraussetzen, dies bedeutet
 $\|\nabla^2 f(x) - \nabla^2 f(\bar{x})\| \leq \Pi \|x - \bar{x}\|$ für $x, \bar{x} \in D$.

Zur Vereinfachung schreiben wir für die Matrix $\nabla^2 f(x^*)$ zukünftig G , und wie gebräuchlich benützen wir $\sigma(x, \bar{x})$ anstelle von $\max(\|x - x^*\|, \|\bar{x} - x^*\|)$. Ohne Beweis zitieren wir aus Broyden, Dennis und Moré [73] die folgende grundlegende Aussage:

Lemma 3.1.1: Seien die Voraussetzungen 1. und 2. erfüllt und $x, \bar{x}$ sowie x^* beliebige Punkte aus D , dann gilt

$$\|\nabla f(\bar{x}) - \nabla f(x) - G(\bar{x} - x)\| \leq \Pi \sigma(x, \bar{x}) \|\bar{x} - x\|. \quad (3.1.1)$$

Der folgende Satz ist eine wohlbekannte Modifikation von Theorem 3.4 aus Broyden, Dennis und Moré [73], dabei steht $\|H\|_M$ für eine beliebige Matrixnorm. Wegen der Äquivalenz der Normen gibt es jedoch eine Konstante η , sodass $\|H\| \leq \eta \|H\|_M$ gilt.

Satz 3.1.2: Seien die Voraussetzungen 1. bis 3. erfüllt. Sei x_k eine Folge von Punkten in D und H_k eine Folge regulärer Matrizen mit der Eigenschaft

$$\|H_{k+1} - G^{-1}\|_M \leq [1 + \alpha_1 \sigma(x_k, x_{k+1})] \|H_k - G^{-1}\|_M + \alpha_2 \sigma(x_k, x_{k+1}) \quad (3.1.2)$$

und sei

$$z_k := H_k^{-1} s_k + g_k . \quad (3.1.3)$$

Dann gibt es zu jedem $r \in (0, 1)$ Zahlen $\epsilon \geq 0$, $\delta \geq 0$, $\beta \geq 0$, sodass aus

$$\|x_0 - x^*\| \leq \epsilon \quad (3.1.4)$$

$$\|H_0 - G^{-1}\|_M \leq \delta \quad (3.1.5)$$

$$\|z_k\| \leq \beta \|g_k\|, \forall k \quad (3.1.6)$$

folgt

$$\|x_{k+1} - x^*\| \leq r \|x_k - x^*\|$$

und $\|H_k\|$ sowie $\|H_k^{-1}\|$ sind gleichmässig beschränkt.

Beweis : Es gilt $\|G\| \leq \Pi$ und $\|G^{-1}\| \leq \pi^{-1}$. Sei $r \in (0, 1)$ vorgegeben und nun wähle man $\epsilon \geq 0$, $\delta \geq 0$ und $\beta \geq 0$, sodass gilt

$$2\eta\delta\Pi + (2\eta\delta + \frac{1}{\pi})\Pi(\epsilon + \beta) \leq r \quad (3.1.7)$$

$$(2\alpha_1\delta + \alpha_2)\frac{\epsilon}{1-r} \leq \delta . \quad (3.1.8)$$

Wegen (3.1.7) ist $2\eta\delta\Pi \leq r < 1$, zusammen mit $\|H_0 - G^{-1}\| \leq \eta\delta < 2\eta\delta$ folgt aus dem Störungslemma (vgl. Ortega und Rheinboldt [70], p 45) $\|H_0^{-1}\| \leq \Pi(1-r)^{-1}$. Zudem gilt $\|H_0\| \leq \|H_0 - G^{-1}\| + \|G^{-1}\| \leq 2\eta\delta + \pi^{-1}$. Aus (3.1.3), der Definition von z_k , erhalten wir

$$x_1 - x^* = -H_0[g_0 - g^* - G(x_0 - x^*)] + (I - H_0G)(x_0 - x^*) + H_0z_0 ,$$

verwenden wir noch Lemma 3.1.1 so folgt

$$\|x_1 - x^*\| \leq \left[(2\eta\delta + \frac{1}{\pi})\Pi\epsilon + 2\eta\delta\Pi + (2\eta\delta + \frac{1}{\pi})\beta\Pi \right] \|x_0 - x^*\|$$

und somit gemäss (3.1.7) die Behauptung für $k=0$.

Der Beweis wird nun mittels Induktion über k geführt; wir setzen deshalb voraus, dass $\|H_i - G^{-1}\|_M \leq 2\delta$ und $\|x_{i+1} - x^*\| \leq r \|x_i - x^*\|$ für $i = 0, 1, \dots, k$ erfüllt ist. Aus (3.1.2) ergibt sich

$$\|H_{i+1} - G^{-1}\|_M - \|H_i - G^{-1}\|_M \leq (2\alpha_1\delta + \alpha_2) r^{i+1} \epsilon$$

und daraus durch Addieren

$$\|H_{k+1} - G^{-1}\|_M \leq \|H_0 - G^{-1}\|_M + (2\alpha_1\delta + \alpha_2) \frac{\epsilon}{1-r},$$

somit wegen (3.1.8) $\|H_{k+1} - G^{-1}\|_M \leq 2\delta$. Wie oben folgt

$$\|H_{k+1}^{-1}\| \leq \frac{\Pi}{1-r}, \quad \|H_{k+1}\| \leq 2\eta\delta + \frac{1}{\pi} \quad \text{und} \quad \|x_{k+2} - x^*\| \leq r \|x_{k+1} - x^*\|.$$

□

Offensichtlich ist die Bedingung (3.1.6) trivialerweise erfüllt, sofern $x_{k+1} = x_k - H_k g_k$ gilt. Die zentrale Voraussetzung, Ungleichung (3.1.2), verlangt nun eine sorgfältige Analyse der verwendeten Update-Formel. Broyden, Dennis und Moré [73] haben dazu die folgende Aussage erarbeitet.

Lemma 3.1.3: Sei M eine reguläre, symmetrische $n \times n$ -Matrix, sodass

$$\|Mw - M^{-1}v\| \leq \beta \|M^{-1}v\| \quad (3.1.9)$$

gilt für $\beta \in [0, 1/3]$ und $v, w \in \mathbb{R}^n$ mit $v \neq 0$. Sei B eine symmetrische Matrix und

$$\bar{B} = B + \frac{(u - Bv)w^T + w(u - Bv)^T}{v^T w} - \frac{v^T(u - Bv)w w^T}{(v^T w)^2}. \quad (3.1.10)$$

Für jede symmetrische Matrix $A \neq B$ und die Norm $\|Q\|_M := \|MQM\|_F$ gibt es Zahlen $\alpha \in (0, 1]$, $\hat{\alpha}_1, \hat{\alpha}_2$, welche nur von M und n abhängig sind, sodass

$$\begin{aligned} \|\bar{B} - A\|_M \leq & \left(\sqrt{1 - \alpha^2 \theta^2} + \hat{\alpha}_1 \frac{\|Mw - M^{-1}v\|}{\|M^{-1}v\|} \right) \|B - A\|_M \\ & + \hat{\alpha}_2 \frac{\|u - Av\|}{\|M^{-1}v\|} \end{aligned} \quad (3.1.11)$$

gilt, wobei θ für $\frac{\|M(B-A)v\|}{\|B-A\|_M \|M^{-1}v\|}$ steht.

Für die folgenden Überlegungen werden wir uns auf die BFGS-Formel festlegen. Die nächste Aussage entspricht Lemma 3.2 in Dennis und Moré [74].

Lemma 3.1.4 : Seien die Voraussetzungen 1. bis 3. erfüllt, dann existieren eine Umgebung U von x^* und Konstanten $\alpha \in (0, 1]$, α_1, α_2 , sodass für jede vorgegebene symmetrische Matrix H , für Punkte $x, \bar{x} \in U$ und für $\bar{H}$ erzeugt nach der BFGS-Formel gilt

$$\|\bar{H} - G\|_M \leq [\sqrt{1 - \alpha\theta^2} + \alpha_1\sigma(x, \bar{x})] \|H - G\|_M + \alpha_2\sigma(x, \bar{x}) \quad , \quad (3.1.12)$$

$$\text{mit } M := G^{\frac{1}{2}} \text{ und } \theta = \frac{\|M(H - G^{-1})y\|}{\|H - G^{-1}\|_M \|M^{-1}y\|} \quad .$$

Beweis : In Lemma 3.1.3 wählen wir für $B, \bar{B}$ und A die Matrizen $H, \bar{H}$ und G^{-1} und setzen $u = w = s$, $v = y$, dadurch wird (3.1.10) zur BFGS-Formel. Der Quotient in (3.1.11) hinter $\hat{\alpha}_1$ ist wegen $\|y\| \leq \|G^{\frac{1}{2}}\| \|G^{-\frac{1}{2}}y\|$ kleiner oder gleich

$$\frac{\|G^{-\frac{1}{2}}\| \|Gs - y\| \|G^{\frac{1}{2}}\|}{\|y\|} \quad ,$$

unter Verwendung von Lemma 3.1.1 und

$$\pi\|s\| \leq \|y\| \quad (3.1.13)$$

kann dieser Term durch $\Pi\pi^{-1}\|G^{-\frac{1}{2}}\|\|G^{\frac{1}{2}}\|\sigma(x, \bar{x})$ nach oben beschränkt werden. Wählen wir nun

$$U := \{x \in D : \|x - x^*\| \leq \pi(3\Pi\|G^{-\frac{1}{2}}\|\|G^{\frac{1}{2}}\|)^{-1}\} \quad ,$$

so ist Voraussetzung (3.1.9) erfüllt. Der Quotient in (3.1.11) hinter $\hat{\alpha}_2$ lässt sich ganz analog nach oben abschätzen. Mit $\alpha_1 = \hat{\alpha}_1\Pi\pi^{-1}\|G^{-\frac{1}{2}}\|\|G^{\frac{1}{2}}\|$ und $\alpha_2 = \hat{\alpha}_2\Pi\pi^{-1}\|G^{\frac{1}{2}}\|$ ist die Behauptung gezeigt. $\square$

Satz 3.1.2 und Lemma 3.1.4 implizieren die lokale und lineare Konvergenz des BFGS-Verfahrens, wenn immer $\|x_0 - x^*\|$ und $\|H_0 - G^{-1}\|$ hinreichend klein sind und sofern $x_{k+1} = x_k - H_k g_k$ gilt (d.h. im Gegensatz zum tatsächlich formulierten Verfahren wird Schrittlänge 1 verwendet).

Will man die eben dargestellte Konvergenztheorie auf die Verfahren aus Kapitel 2 übertragen, so ergeben sich zwei kritische Stellen in der Beweisführung. Da bei den konischen Näherungsfunktionen die Schrittlänge, bei der rationalen Näherungsfunktion gar die Suchrichtung modifiziert wird, muss die Ungleichung (3.1.6) aus Satz 3.1.2 explizit bewiesen werden. Zudem entspricht der Vektor y nicht mehr der Differenz der

Gradienten, sodass in Lemma 3.1.4 das Lemma 3.1.1 nicht mehr verwendet werden kann. Für die Ungleichung (3.1.13) erhält man leicht einen Ersatz, sofern man $y^T s$ nach unten durch $\|s\|^2$ abschätzen kann. Im folgenden werden die notwendigen Modifikationen untersucht.

3.2 Modifikationen im Beweis bei Verwendung konischer Näherungsfunktionen

Zuerst befassen wir uns mit dem Verhalten der Grössen γ und ρ in der Nähe eines lokalen Minimums. Zu diesem Zwecke fügen wir den Voraussetzungen aus Kapitel 3.1 eine weitere hinzu:

4. Die Umgebung D von x^* werde derart eingeschränkt, dass für $x \in D$ gilt

$$\|\nabla f(x)\| \leq \frac{3\pi^2}{8\Pi} .$$

Lemma 3.2.1: Seien die Voraussetzungen 1. bis 4. erfüllt, dann gelten die Ungleichungen

$$\rho \geq \frac{\pi}{4} \|s\|^2 \quad (3.2.1)$$

$$\left| \frac{1}{\gamma} - 1 \right| \leq \frac{2\Pi}{3\pi} \|s\| \quad (3.2.2)$$

sofern $x, \bar{x} \in D$.

Beweis: Aus der Gleichung

$$\rho^2 = (\bar{f} - f - g^T s)^2 + g^T s [\bar{f} - f - g^T s - (f - \bar{f} + \bar{g}^T s)] \quad (3.2.3)$$

erhalten wir aufgrund des Satzes von Taylor

$$\rho^2 \geq \left[\frac{1}{2} s^T \nabla^2 f(u) s \right]^2 - |g^T s| \cdot \left| \frac{1}{2} s^T [\nabla^2 f(u) - \nabla^2 f(v)] s \right|$$

für zwei Vektoren $u, v \in D$ mit $u - v$ parallel zu s . Wegen den Voraussetzungen 1. bis 3. folgt $\rho^2 \geq \frac{1}{4} [\pi^2 - 2\Pi\|g\|] \|s\|^4$ und dank der Voraussetzung 4. ergibt sich die Behauptung (3.2.1).

Gemäss (2.1.12) gilt

$$\frac{1}{\gamma} - 1 = \frac{\bar{f} - f - g^T s - \rho}{g^T s} ,$$

durch Erweitern mit $\bar{f} - f - g^T s + \rho$ wird dieser Ausdruck zu

$$\frac{\bar{f} - f - g^T s - (f - \bar{f} + \bar{g}^T s)}{\bar{f} - f - g^T s + \rho} = \frac{\frac{1}{2} s^T [\nabla^2 f(u) - \nabla^2 f(v)] s}{\frac{1}{2} s^T \nabla^2 f(v) s + \rho}, \quad (3.2.4)$$

wobei (3.2.3) zur Anwendung kommt und wie oben $u, v \in D$ gilt. Aufgrund unserer Voraussetzungen kann der Betrag des Zählers durch $\frac{1}{2} \Pi \|s\|^3$ nach oben und der Nenner durch $\frac{1}{4} \Pi \|s\|^2$ nach unten abgeschätzt werden, womit das Lemma bewiesen ist. $\square$

Lemma 3.2.2: Unter den Voraussetzungen 1. bis 4. existiert eine positive reelle Konstante τ mit der Eigenschaft

$$\frac{1}{\tau} \leq \gamma \leq \tau. \quad (3.2.5)$$

Beweis: Für den Zähler in (3.2.4) gilt die Abschätzung

$$s^T \nabla^2 f(u) s - s^T \nabla^2 f(v) s \leq \max [s^T \nabla^2 f(u) s, s^T \nabla^2 f(v) s] \leq \Pi \|s\|^2,$$

womit wir $\frac{1}{\gamma} - 1 \leq \frac{2\Pi}{3\pi}$ erhalten und deshalb

$$\tau = \frac{2\Pi}{3\pi} + 1$$

wählen. Aus (2.1.12) folgt

$$\gamma - 1 = \frac{f - \bar{f} + \bar{g}^T s - \rho}{-\bar{g}^T s}$$

und unter Beachtung von $\rho^2 = (f - \bar{f} + \bar{g}^T s)^2 - \bar{g}^T s [f - \bar{f} + \bar{g}^T s - (\bar{f} - f - g^T s)]$ ergibt sich

$$\gamma - 1 = \frac{\frac{1}{2} s^T [\nabla^2 f(u) - \nabla^2 f(v)] s}{\frac{1}{2} s^T \nabla^2 f(v) s + \rho}.$$

Analog zum obigen Fall folgt die rechte Seite unserer Aussage. $\square$

Nun können wir die entsprechende Aussage zu Satz 3.1.2 formulieren:

Satz 3.2.3: Seien die Voraussetzungen 1. bis 4. erfüllt. Sei x_k eine Folge von Punkten in D , H_k eine Folge regulärer Matrizen mit der Eigenschaft (3.1.2) und sei a_k eine Folge von Vektoren aus $\mathbb{R}^n$, gewählt gemäss den Varianten *con1* bis *con6*

derart, dass gilt

$$x_{k+1} = x_k - \frac{1}{1 - a_k^T H_k g_k} H_k g_k . \quad (3.2.6)$$

Dann gibt es zu jedem $r \in (0, 1)$ Zahlen $\epsilon \geq 0$, $\delta \geq 0$, $\beta \in [0, 1]$, sodass aus

$$\|x_0 - x^*\| \leq \epsilon \quad (3.2.7)$$

$$\|H_0 - G^{-1}\|_M \leq \delta \quad (3.2.8)$$

$$|a_0^T H_0 g_0| \leq \frac{1}{2} \beta \quad (3.2.9)$$

folgt

$$\|x_{k+1} - x^*\| \leq r \|x_k - x^*\|$$

und $\|H_k\|$ sowie $\|H_k^{-1}\|$ sind gleichmässig beschränkt.

Beweis : Der Unterschied zu Satz 3.1.2 besteht darin, dass wir hier

$$z_k = \frac{a_k^T H_k g_k}{1 - a_k^T H_k g_k} g_k$$

verwenden. Der Beweis verläuft analog zu jenem von Satz 3.1.2, anstelle von (3.1.6) steht jedoch nur (3.2.9), weshalb wir zusätzlich die Ungleichung

$$\left| \frac{a_k^T H_k g_k}{1 - a_k^T H_k g_k} \right| \leq \beta \quad (3.2.10)$$

untersuchen müssen. Aus $|a_k^T H_k g_k| \leq \frac{1}{2}$ erhält man $\frac{1}{2} \leq |1 - a_k^T H_k g_k|^{-1} \leq 2$. Wählt man nun $\beta \leq 1$ und beweist $|a_k^T H_k g_k| \leq \frac{1}{2} \beta$ für $k \geq 1$, so ergibt sich unmittelbar die gewünschte Ungleichung (3.2.10). Im folgenden werden wir für jede Variante aus Kapitel 2 die Grösse $|a_{k+1}^T H_{k+1} g_{k+1}|$ durch einen Ausdruck der Form $\psi \cdot \|x_k - x^*\|$ nach oben abschätzen. Dies soll innerhalb des Induktionsschrittes geschehen, womit wir für $i = 1, \dots, k+1$

$$\|H_i^{-1}\| \leq \frac{\Pi}{1-r}, \quad \|H_i\| \leq 2\eta\delta + \frac{1}{\pi} \quad \text{und} \quad \|x_i - x^*\| \leq r \|x_{i-1} - x^*\|$$

voraussetzen dürfen. Da nun $\|x_k - x^*\| \leq \epsilon$ gilt, setzen wir $\beta = 2\psi\epsilon$ und können durch

die Wahl von ϵ erreichen, dass $\beta \leq 1$ und die Ungleichung 3.1.7 erfüllt ist. Im folgenden unterdrücken wir in der gewohnten Weise die Indizes k und $k+1$.

Bei der Variante *con1* gilt

$$\|\bar{a}\| \leq \left| \frac{1}{\gamma} - 1 \right| \frac{1}{\|s\|} \leq \frac{2\Pi}{3\pi},$$

deshalb erhalten wir

$$|\bar{a}^T \bar{H} \bar{g}| \leq \frac{2\Pi^2}{3\pi} \left(2\eta\delta + \frac{1}{\pi} \right) \|x - x^*\|,$$

und wählen

$$\psi = \frac{2\Pi^2}{3\pi} \left(2\eta\delta + \frac{1}{\pi} \right).$$

Für *con3* ist

$$\|\bar{a}\| \leq \left| \frac{1}{\gamma} - 1 \right| \frac{\|\bar{g} - g\| + \left| \frac{1}{\gamma} - 1 \right| \|g\|}{|\bar{g}^T s + f - \bar{f} + \rho|} \leq \frac{8\Pi^2 + 2\Pi\pi}{9\pi^2} \quad (3.2.11)$$

und darum gilt

$$|\bar{a}^T \bar{H} \bar{g}| \leq \Pi^2 \frac{8\Pi + 2\pi}{9\pi^2} \left(2\eta\delta + \frac{1}{\pi} \right) \|x - x^*\|.$$

Für die verbleibenden Varianten benötigen wir einige zusätzliche Ungleichungen. Wegen

$$g^T H g \geq (x - x^*)^T \nabla^2 f(u)^T H \nabla^2 f(u) (x - x^*) \geq \frac{\|x - x^*\|^2}{\|\nabla^2 f(u)^{-T} H^{-1} \nabla^2 f(u)^{-1}\|}$$

ergibt sich

$$\frac{1}{g^T H g} \leq \frac{\Pi}{(1-r)\pi^2 \|x - x^*\|^2}$$

und

$$\left| \frac{\bar{g}^T s}{g^T s} \right| \quad \text{sowie} \quad \frac{\bar{g}^T \bar{H} \bar{g}}{g^T H g} \leq \frac{\Pi^3}{(1-r)\pi^2} \left(2\eta\delta + \frac{1}{\pi} \right).$$

Unter Beachtung von $|g^T s| = \left| \frac{g^T H g}{1 - a^T H g} \right| \geq \frac{1}{2} g^T H g$ folgt zudem

$$\left| \frac{g^T \bar{H} \bar{g}}{g^T s} \right| \quad \text{ sowie } \quad \left| \frac{\bar{g}^T \bar{H} \bar{g}}{g^T s} \right| \leq \frac{2\Pi^3}{(1-r)\pi^2} \left(2\eta\delta + \frac{1}{\pi} \right).$$

Für die Variante *con2* gilt

$$|\bar{a}^T \bar{H} \bar{g}| \leq \left| \frac{1}{\gamma} - 1 \right| \cdot \left| \frac{g^T \bar{H} \bar{g}}{g^T s} \right| \leq \frac{4\Pi^4}{3(1-r)\pi^3} \left(2\eta\delta + \frac{1}{\pi} \right) \|s\|,$$

wegen $\|s\| \leq 2\|x - x^*\|$ wählen wir deshalb

$$\psi = \frac{8\Pi^4}{3(1-r)\pi^3} \left(2\eta\delta + \frac{1}{\pi} \right).$$

Bei der Variante *con4* ist

$$\bar{a} = \left(\frac{1}{\gamma} - 1 \right) \left[g \left(1 + \gamma \frac{\bar{g}^T s}{g^T s} \right) \frac{s}{s^T s} - \frac{\gamma}{g^T s} \bar{g} \right],$$

deshalb heisst die entsprechende Ungleichung

$$|\bar{a}^T \bar{H} \bar{g}| \leq \left| \frac{1}{\gamma} - 1 \right| \cdot \left| 1 + \gamma \frac{\bar{g}^T s}{g^T s} \right| \frac{\|\bar{H}\| \cdot \|\bar{g}\|}{\|s\|} + \left| \frac{1}{\gamma} - 1 \right| \cdot \left| \gamma \frac{\bar{g}^T \bar{H} \bar{g}}{g^T s} \right|$$

und setzen wir

$$\psi = \frac{2\Pi^2}{3\pi} \left(2\eta\delta + \frac{1}{\pi} \right) \left[1 + \frac{\tau\Pi^2}{(1-r)\pi^2} \left(2\eta\delta\Pi + \frac{\Pi}{\pi} + 4 \right) \right].$$

Ganz ähnlich gilt bei Variante *con5*

$$\bar{a} = \left(\frac{1}{\gamma} - 1 \right) \left[\left(1 + \gamma \frac{\bar{g}^T s}{g^T s} \right) \frac{g}{g^T s} - \frac{\gamma}{g^T s} \bar{g} \right]$$

$$|\bar{a}^T \bar{H} \bar{g}| \leq \left| \frac{1}{\gamma} - 1 \right| \cdot \left[\left| 1 + \gamma \frac{\bar{g}^T s}{g^T s} \right| \cdot \left| \frac{g^T \bar{H} g}{g^T s} \right| + \left| \gamma \frac{\bar{g}^T \bar{H} \bar{g}}{g^T s} \right| \right]$$

und wir schreiben

$$\psi = \frac{8\Pi^4}{3\pi^3(1-r)} \left(2\eta\delta + \frac{1}{\pi} \right) \left[1 + \tau + \frac{\tau\Pi^3}{(1-r)\pi^2} \left(2\eta\delta + \frac{1}{\pi} \right) \right].$$

Für die verbleibende Variante *con6* kann wegen

$$\bar{a} = \left(\frac{1}{\gamma} - 1\right) \left[\left(1 + \gamma \frac{\bar{g}^T s}{g^T s}\right) \frac{\bar{g} - g}{\bar{g}^T s - g^T s} - \frac{\gamma}{g^T s} \bar{g} \right]$$

die Grösse $\left| \bar{a}^T \bar{H} \bar{g} \right|$ nach oben durch

$$\left| \frac{1}{\gamma} - 1 \right| \cdot \left| 1 + \gamma \frac{\bar{g}^T s}{g^T s} \right| \frac{\|\bar{g} - g\| \|\bar{H}\|}{\left| \bar{g}^T s - g^T s \right|} \|\bar{g}\| + \left| \frac{1}{\gamma} - 1 \right| \cdot \left| \gamma \frac{\bar{g}^T \bar{H} \bar{g}}{g^T s} \right|$$

abgeschätzt werden, weshalb wir

$$\psi = \frac{2\Pi^3}{3\pi^2} \left(2\eta\delta + \frac{1}{\pi} \right) \left[1 + \frac{\tau\Pi^3}{\pi^2(1-r)} \left(2\eta\delta + \frac{1}{\pi} \right) + \frac{2\tau\Pi}{\pi(1-r)} \right]$$

setzen können. Damit ist die Behauptung gezeigt. $\square$

Bei konischen Näherungsfunktionen kann die BFGS-Formel gemäss der Gleichung (2.1.9) in die folgende Form

$$\bar{H} = \gamma^2 \left(I - \frac{s y^T}{y^T s} \right) H \left(I - \frac{y s^T}{y^T s} \right) + \frac{s s^T}{y^T s}.$$

gebracht werden, wie dies von Sorensen [80] dargestellt wurde. Dies ist nicht die einzige Möglichkeit. Wir werden mit

$$\bar{H} = \left(I - \frac{s y^T}{y^T s} \right) H \left(I - \frac{y s^T}{y^T s} \right) + \gamma^2 \frac{s s^T}{y^T s} \quad (3.2.12)$$

arbeiten, wodurch die Kovergenzzuntersuchung deutlich vereinfacht wird. So haben wir nur noch die nachstehende Erweiterung von Lemma 3.1.1 zu beweisen.

Lemma 3.2.4 : Unter den Voraussetzungen 1. bis 4. existiert eine reelle Konstante κ , sodass für $x, \bar{x} \in D$ und für alle Varianten mit Ausnahme von *con5* gilt

$$\left\| \frac{1}{\gamma^2} y - G s \right\| \leq \kappa \sigma(x, \bar{x}) \|s\|. \quad (3.2.13)$$

Beweis : Nach der Definition von y ist $\| \gamma^{-2} y - G s \|$ nach oben beschränkt durch

$$\|(\frac{1}{\gamma}-1)\bar{g}-(\frac{1}{\gamma^2}-1)g\|+\|\frac{g^{Ts}}{\gamma^2}\bar{a}\|+\|\bar{g}-g-Gs\|.$$

Für den hintersten Summanden gilt eine Ungleichung vom gewünschten Typ, gemäss Lemma 3.1.1. Der erste Summand lässt sich nach Ausklammern von $|\frac{1}{\gamma}-1|$ durch

$$\frac{2\Pi^2}{3\pi}(\tau+2)\sigma(x,\bar{x})\|s\|$$

nach oben abschätzen. Somit bleibt zu zeigen, dass eine entsprechende Ungleichung auch für den mittleren Summanden gilt.

Bei den Varianten *con1* und *con2* gilt

$$\frac{1}{\gamma^2}\|g^{Ts}\bar{a}\|\leq\tau^2|\frac{1}{\gamma}-1|\|g\|\leq\frac{2\Pi^2\tau^2}{3\pi}\sigma(x,\bar{x})\|s\|.$$

Für Variante *con3* folgt aus (3.2.11)

$$\frac{1}{\gamma^2}\|g^{Ts}\bar{a}\|\leq\frac{2\Pi^2\tau^2}{9\pi^2}(4\Pi+\pi)\sigma(x,\bar{x})\|s\|.$$

Bei Variante *con4* lautet die entsprechende Ungleichung

$$\frac{1}{\gamma^2}\|g^{Ts}\bar{a}\|\leq\frac{1}{\gamma}|\frac{1}{\gamma}-1|\left(\frac{1}{\gamma}\|g\|+2\|\bar{g}\|\right),$$

sodass man die rechte Seite durch

$$\frac{2\Pi^2\tau}{3\pi}(2+\tau)\sigma(x,\bar{x})\|s\|$$

nach oben abschätzen kann.

In der Variante *con6* gilt

$$\frac{1}{\gamma^2}\|g^{Ts}\bar{a}\|\leq\frac{1}{\gamma}|\frac{1}{\gamma}-1|\left(\left|\frac{1}{\gamma}g^{Ts}+\bar{g}^{Ts}\right|\frac{\Pi}{\pi\|s\|}+\|\bar{g}\|\right)$$

und für die rechte Seite erhält man die folgende obere Schranke

$$\frac{2\Pi^2\tau}{3\pi^2}\left[(\tau+1)\Pi+\pi\right]\sigma(x,\bar{x})\|s\|.$$

□

Lemma 3.2.4 wird benötigt um die Aussage von Lemma 3.1.4 zu garantieren und damit um die Gültigkeit von Ungleichung (3.1.2) nachzuweisen. Integriert man nun die beiden Lemmata in den Beweis von Satz 3.1.2, so erhält man innerhalb des Induktionsschrittes die Beschränktheit von $1 + \gamma \frac{\bar{g}^{Ts}}{g^{Ts}}$. Deshalb ist die Ausnahme, welche Variante *con5* bei

Lemma 3.2.4 macht, nicht gravierend.

Zusammenfassend halten wir fest, dass die BFGS-Verfahren mit konischer Näherungsfunktion lokal und linear konvergieren.

3.3 Modifikationen im Beweis bei Verwendung rationaler Näherungsfunktionen

Zuerst untersuchen wir wieder das lokale Verhalten von γ und y^{Ts} .

Lemma 3.3.1: Unter den Voraussetzungen 1. bis 3. gelten die beiden folgenden Ungleichungen

$$\frac{1}{\tau} \leq \gamma \leq \tau, \quad \text{mit } \tau = \frac{\Pi}{\pi}, \quad (3.3.1)$$

$$\left| \frac{1}{\gamma} - 1 \right| \leq \frac{\Pi}{\pi} \|s\|. \quad (3.3.2)$$

Mit der zusätzlichen Voraussetzungen 4. ergibt sich zudem

$$y^{Ts} \geq \frac{\pi^2}{4\Pi} \|s\|^2. \quad (3.3.3)$$

Beweis: Aus (2.2.14) und dem Satz von Taylor folgen unmittelbar die ersten beiden Behauptungen. Für $\frac{1}{2}y^{Ts}$ kann gemäss (2.2.16) der Ausdruck

$$\frac{(\bar{f} - f - g^{Ts})^2 + g^{Ts} [\bar{f} - f - g^{Ts} - (f - \bar{f} + \bar{g}^{Ts})]}{\bar{f} - f - g^{Ts}}$$

geschrieben werden, daraus erhält man wie im Beweis von Lemma 3.2.1 die letzte Behauptung. □

Als Analogon zu Satz 3.1.2 erhalten wir

Satz 3.3.2: Seien die Voraussetzungen 1. bis 3. erfüllt. Sei x_k eine Folge von Punkten in D , H_k eine Folge regulärer Matrizen mit der Eigenschaft (3.1.2) und sei a_k eine Folge von Vektoren in $\mathbb{R}^n$, sodass gilt

$$x_{k+1} = x_k - H_k g_k + \mu_k H_k a_k \quad (3.3.4)$$

$$a_{k+1} = \frac{\frac{1}{\gamma_k} - 1}{g_{k+1}^T s - g_k^T s} (g_{k+1} - g_k) . \quad (3.3.5)$$

Dann gibt es zu jedem $r \in (0, 1)$ Zahlen $\epsilon \geq 0$, $\delta \geq 0$, $\beta \geq 0$, sodass aus

$$\|x_0 - x^*\| \leq \epsilon \quad (3.3.6)$$

$$\|H_0 - G^{-1}\|_M \leq \delta \quad (3.3.7)$$

$$\|\mu_0 a_0\| \leq \beta \|g_0\| \quad (3.3.8)$$

folgt

$$\|x_{k+1} - x^*\| \leq r \|x_k - x^*\|$$

und $\|H_k\|$ sowie $\|H_k^{-1}\|$ sind gleichmässig beschränkt.

Beweis : Der Beweis verläuft wieder analog zum Beweis von Satz 3.1.2, allerdings müssen wir innerhalb des Induktionsschrittes für $z_{k+1} = \mu_{k+1} a_{k+1}$ die Ungleichung (3.1.6) nachweisen. Wir dürfen voraussetzen, dass für $i=1, \dots, k+1$

$$\|H_i^{-1}\| \leq \frac{\Pi}{1-r}, \quad \|H_i\| \leq 2\eta\delta + \frac{1}{\pi} \quad \text{und} \quad \|x_i - x^*\| \leq r \|x_{i-1} - x^*\|$$

gilt. Im folgenden unterdrücken wir wie gewohnt die Indizes. Aus

$$\bar{\mu} = \frac{\bar{p}^2}{\bar{\sigma} + \sqrt{\bar{\sigma}^2 - \bar{p}^2}}$$

folgt wegen $\bar{\sigma} > 0$

$$\bar{\mu} \leq \frac{\bar{p}^2}{\sqrt{\bar{\sigma}^2 - \bar{p}^2}} .$$

Im untenstehenden Ausdruck liefert (2.2.18) die Gleichung, während durch Lemma 3.3.1 das Ungleichheitszeichen impliziert wird:

$$\bar{\sigma}^2 - \bar{p}^2 = \frac{2(f - \bar{f} + \bar{g}^T s)}{\gamma (\bar{a}^T s)^2 \bar{a}^T \bar{H} \bar{a}} \geq \frac{\pi^4}{\Pi^3 \bar{a}^T \bar{H} \bar{a}} .$$

Unter Verwendung von (2.2.6) erhalten wir

$$\|\bar{\mu} \bar{a}\| \leq \sqrt{\frac{\Pi^3}{\bar{a}^T \bar{H} \bar{a}}} \cdot \frac{\bar{g}^T \bar{H} \bar{g}}{\pi^2} \|\bar{a}\|.$$

Beachten wir zum einen die Ungleichung $\|\bar{a}\|^2 \|\bar{H}^{-1}\|^{-1} \leq \bar{a}^T \bar{H} \bar{a}$ und zum andern die Induktionsvoraussetzungen, so folgt

$$\|\bar{z}\| \leq \frac{\Pi^3 (2\eta\delta + \pi^{-1})}{\pi^2 \sqrt{1-r}} \|\bar{x} - x^*\| \cdot \|\bar{g}\|. \quad (3.3.9)$$

Wenn wir nun für β den Ausdruck

$$\frac{\Pi^3 (2\eta\delta + \pi^{-1})}{\pi^2 \sqrt{1-r}} \epsilon$$

setzen, erhalten wir die gewünschte Behauptung. $\square$

Als letzte Aussage benötigen wir noch eine Modifikation von Lemma 3.1.1.

Lemma 3.3.3 : Es seien die Voraussetzungen 1. bis 3. erfüllt, dann existiert eine reelle Konstante $\kappa > 0$, sodass für $x, \bar{x} \in D$ gilt

$$\|y - Gs\| \leq \kappa \sigma(x, \bar{x}) \|s\|.$$

Beweis : Gemäss der Definition von y und der Wahl von a gilt

$$y - Gs = \bar{g} - g - Gs - \left(\frac{1}{\gamma} - 1\right)g + \left(\frac{1}{\gamma} - 1\right) \frac{f - \bar{f}}{\bar{g}^T s - g^T s} (\bar{g} - g).$$

Nach den Lemmata 3.1.1 und 3.3.1 und dem Satz von Taylor folgt

$$\|y - Gs\| \leq \Pi \sigma(x, \bar{x}) \|s\| + \frac{\Pi}{\pi} \|s\| \left[\Pi \|x - x^*\| + \frac{\|\nabla f(u)\| \Pi}{\pi} \right],$$

wobei $u \in [x, \bar{x}]$ gilt. Wegen $\|\nabla f(u)\| \leq \Pi \|u - x^*\| \leq \Pi \sigma(x, \bar{x})$ ist die Behauptung für

$$\kappa = \Pi + \frac{\Pi^2}{\pi} + \frac{\Pi^3}{\pi^2}$$

erfüllt. $\square$

Es ist hiermit gezeigt, dass das BFGS-Verfahren mit rationaler Näherungsfunktion lokal und linear konvergiert.

3.4 Superlineare Konvergenz

Sofern die Voraussetzungen 1. bis 3. erfüllt und $\|x_0 - x^*\|$ sowie $\|H_0 - G^{-1}\|$ hinreichend klein sind, so folgt aus Lemma 3.1.4 und Satz 3.1.2

$$\lim_{k \rightarrow \infty} \frac{\|(H_k - G^{-1})y_k\|}{\|y_k\|} = 0, \quad (3.4.1)$$

wie dies von Dennis und Moré [74] gezeigt worden ist. Mit diesem Ergebnis lässt sich die superlineare Konvergenz der von uns betrachteten Verfahren nachweisen.

Lemma 3.4.1: Es seien die Voraussetzungen 1. bis 3. erfüllt, x_k eine gegen x^* konvergente Folge und $\|H_k\|$ sowie $\|H_k^{-1}\|$ gleichmässig beschränkt. Falls für $k \rightarrow \infty$ gilt

$$\frac{\|(H_k - G^{-1})y_k\|}{\|y_k\|} \rightarrow 0 \quad (3.4.2)$$

$$\frac{\|y_k - g_{k+1} + g_k\|}{\|s_k\|} \rightarrow 0 \quad (3.4.3)$$

$$\frac{\|y_k - G s_k\|}{\|s_k\|} \rightarrow 0 \quad (3.4.4)$$

$$\frac{\|z_k\|}{\|g_k\|} \rightarrow 0, \quad (3.4.5)$$

so konvergiert x_k superlinear gegen x^* .

Beweis: Aus der Definition (3.1.3) erhält man

$$(H_k - G^{-1})y_k = H_k(y_k - g_{k+1} + g_k) + H_k g_{k+1} - G^{-1}(y_k - G s_k) - H_k z_k.$$

Anwenden der Dreiecksungleichung und Division durch $\|s_k\|$ liefert

$$\begin{aligned} \frac{\| (H_k - G^{-1}) y_k \|}{\| s_k \|} + \| H_k \| \frac{\| y_k - g_{k+1} + g_k \|}{\| s_k \|} + \| G^{-1} \| \frac{\| y_k - G s_k \|}{\| s_k \|} \\ + \| H_k \| \frac{\| z_k \|}{\| s_k \|} \geq \frac{1}{\| H_k^{-1} \|} \frac{\| g_{k+1} \|}{\| s_k \|} . \end{aligned} \quad (3.4.6)$$

Der zweite und der dritte Summand streben nach den Voraussetzungen gegen Null. Gemäss (3.4.4) existiert eine positive Konstante ξ mit der Eigenschaft $\| y_k - G s_k \| \leq \xi \| s_k \|$. Daher gilt $\| y_k \| \leq \| y_k - G s_k \| + \| G s_k \| \leq (\xi + \Pi) \| s_k \|$, woraus zusammen mit (3.4.2) folgt, dass der erste Summand aus (3.4.6) gegen Null strebt. Wegen (3.4.5) gibt es zu einer positiven Konstanten $\xi < 1$ eine Nummer k_0 , sodass aus $k \geq k_0$ folgt $\| z_k \| \leq \xi \| g_k \|$. Unter Verwendung von (3.1.3) erhält man $\xi \| g_k \| + \| H_k^{-1} \| \| s_k \| \geq \| g_k \|$ und somit

$$\frac{\| H_k^{-1} \|}{1 - \xi} \cdot \frac{1}{\| g_k \|} \geq \frac{1}{\| s_k \|}$$

für $k \geq k_0$. Dies zeigt uns, dass auch der vierte Summand in (3.4.6) gegen Null strebt, als Konsequenz ergibt sich

$$\frac{\| g_{k+1} \|}{\| s_k \|} \rightarrow 0 \quad \text{für} \quad k \rightarrow \infty . \quad (3.4.7)$$

Aus $\| g_{k+1} \| \geq \pi \| x_{k+1} - x^* \|$ und $\| s_k \| \leq \| x_{k+1} - x^* \| + \| x_k - x^* \|$ folgt

$$\frac{\| g_{k+1} \|}{\| s_k \|} \geq \pi \frac{\xi_k}{\xi_k + 1} , \quad \text{mit} \quad \xi_k := \frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} ,$$

sodass wegen (3.4.7) $\xi_k \rightarrow 0$ gelten muss. □

Bei der gewöhnlichen Variable-Metrik-Methode sind (3.4.3) und (3.4.5) gegenstandslos, da die Zähler verschwinden. Die Bedingung (3.4.4) wird durch Lemma 3.1.1 garantiert. Aber auch bei den Verfahren mit konischer beziehungsweise rationaler Näherungsfunktion haben wir keine weiteren Beweise zu führen. Lemma 3.2.4 bzw. Lemma 3.3.3 liefern uns die Voraussetzung (3.4.4). Im Beweis dieser beiden Lemmata ist der Nachweis von (3.4.3) enthalten. Die Gültigkeit von (3.4.5) haben wir in den Beweisen von Satz 3.2.3 und Satz 3.3.2 gezeigt.

Alle hier betrachteten Verfahren sind somit lokal und superlinear konvergent.

4. EMPIRISCHE UNTERSUCHUNG

4.1 Einleitung

Im vorangehenden Kapitel wurde gezeigt, dass die Verfahren aus Kapitel 2 lokal und superlinear konvergieren. Dasselbe gilt allerdings schon für die gewöhnliche Variable-Metrik-Methode. Dies ist nicht überraschend, denn bei lokalen Konvergenzuntersuchungen wird nur eine hinreichend kleine Umgebung des Minimums betrachtet und dort ist jede Funktion durch eine quadratische Funktion approximierbar. Der Vorteil von nichtquadratischen Näherungsfunktionen wird sich deshalb kaum in der Nähe des Optimums zeigen. Liegt der Startpunkt jedoch entfernt, so muss das Verfahren grosse Schritte in Gegenden zurücklegen, wo die Funktion kein quadratisches Verhalten zeigt. Dort können nichtquadratische Näherungsfunktionen zusätzliche Informationen sammeln. Diese Aussage soll im folgenden durch numerische Tests überprüft werden.

4.2 Implementation der Verfahren

Zuerst wurde für die Variable-Metrik-Methode ein Programm in Form einer Fortran77-Subroutine geschrieben, welches die unfaktorierte BFGS-Formel verwendet. Die Arbeiten von Shanno und Phua [78b] sowie von Grandinetti [80] rechtfertigen den Verzicht auf die LDL^T -Darstellung. Der Line search enthält eine kubische Interpolation, wie sie zuerst von Davidon [59] angegeben wurde und seither eine weite Verbreitung gefunden hat (vgl. Fletcher und Powell [63]). Eingebaut ist zudem eine Sicherung, damit das betrachtete Intervall in jedem Durchgang um mindestens $1/20$ verkürzt wird. Bei den ersten n Iterationen wird der Ausgangswert für die Schrittlänge nach der Formel

$$\lambda = \min \left\{ 1, \frac{-2\Delta f}{g^Ts} \right\},$$

ermittelt, nachher setzen wir $\lambda = 1$, denn in der Matrix sollte sich nun genügend Information über die Funktion akkumuliert haben. Obige Formel kann aus (1.2.5) hergeleitet werden und lässt sich beispielsweise bei Fletcher und Powell [63] finden. Der Line search wird abgebrochen, sobald die Bedingungen

$$\bar{f} - f \leq \alpha g^Ts$$

$$\bar{g}^Ts \geq \beta g^Ts$$

erfüllt sind (vgl. Dennis und Moré [74] oder Fletcher [80]). Als konkrete Werte wählen wir

$\alpha = 0.01$ und $\beta = 0.9$. Bei Problemen von mittlerer oder hoher Dimension ist die Skalierung der Ausgangsmatrix für das numerische Verhalten und die Stabilität des Verfahrens ausserordentlich wichtig. Dabei wird die Matrix vor dem ersten Update-Prozess gemäss der Formel

$$\hat{H}_0 = \frac{y_0^T s_0}{y_0^T H_0 s_0} H_0 \quad (4.2.1)$$

transformiert, wie dies von Shanno und Phua [78a] beschrieben wurde.

Zwei Kopien dieses Programms wurden mit den nötigen Änderungen für konische beziehungsweise rationale Näherungsfunktionen versehen. Es handelt sich dabei um eine zusätzliche Bedingung im Line search, vgl. (2.1.11) und (2.2.15), die Erzeugung des Vektors y , vgl. (2.1.9) und (2.2.13), und die Bestimmung der Suchrichtung, vgl. (2.1.4) und (2.2.3). Im besten Fall sind dafür ungefähr zwanzig zusätzliche Programmzeilen nötig.

Da wir die Verfahren anhand bekannter Beispiele testen wollen, verwenden wir die Abbruchbedingung

$$f(x_k) \leq f(x^*) + 10^{-5} \max \{ f(x^*), 10^{-10} \}, \quad (4.2.2)$$

welche praktisch identisch ist mit dem Assessmentkriterium von Gill und Murray [79].

```

C*****
C      SUBROUTINE  UOPTQU
C      -----
C      DIESE SUBROUTINE LOEST FREIE MINIMIERUNGSPROBLEME
C      VERWENDETE METHODE: QUASI-NEWTON MIT QUADRATISCHEM MODELL
C      STANDARDMETHODE
C      VERSION:  16.08.84
C*****
C      SUBROUTINE  UOPTQU(N,X,XX,G,GG,Y,P,H,EPS,FF,ITNU,IEV,KTR,BG)
C      IMPLICIT REAL*8 (A-H,O-Z)
C      DIMENSION X(N),XX(N),G(N),GG(N),Y(N),P(N),H(*)
C
C      VORBEREITUNG DES ERSTEN SCHRITTES
C      -----
C      IEVMAX=IEV
C      CALL FUN(N,X,F,G)
C      DF=DABS(F-FF)
C      IEV=1
C      ITNU=0
C      CALL HMULT(N,H,G,P)
C      GS=0.D0
C      BG=0.D0
C      DO 102 I=1,N
C         GS=GS+G(I)*P(I)
C         BG=DMAX1(BG,DABS(G(I)))
102  CONTINUE
C      KTR=1
C      IF (BG.LT.1.D-30)RETURN
101  CONTINUE
C      KTR=3
C      IF (GS.GE.0.D0)RETURN
C
C      LINE SEARCH
C      -----
C      SL=-2.D0*DF/GS
C      IF ((SL.GT.1.D0).OR.(ITNU.GT.N))SL=1.D0
C      SLU=0.D0
C      FU=F
C      GSU=GS
C      SLO=0.D0
201  DO 202 I=1,N
C         XX(I)=X(I)+SL*P(I)
202  CONTINUE
C      CALL FUN(N,XX,FF,GG)
C      IEV=IEV+1
C      KTR=2
C      IF (IEV.GE.IEVMAX)RETURN
C      GGS=0.D0
C      DO 203 I=1,N
C         GGS=GGS+GG(I)*P(I)
203  CONTINUE
C      DF=F-FF
C      IF (DF.LT.-1.D-2*GS*SL)GOTO 204
C      IF (GGS.LT.0.9D0*GS)GOTO205
C      GOTO 208
204  SLO=SL
C      FO=FF
C      GSO=GGS
C      GOTO 206

```

```

205 IF(SLO.EQ.0.DO)GOTO 207
    SLU=SL
    FU=FF
    GSU=GGG
206 DSL=SLO-SLU
    Z=3.DO*(FU-FO)/DSL+GSU+GSO
    W=DSQRT(Z*Z-GSU*GSO)
    SL=DSL*(1.DO-(GSO+W-Z)/(GSO-GSU+2.DO*W))
    IF(SL.LT.0.05DO*DSL)SL=0.05DO*DSL
    IF(SL.GT.0.95DO*DSL)SL=0.95DO*DSL
    SL=SLU+SL
    GOTO 201
207 DSL=SL-SLU
    SLU=SL
    SL=2.DO*DSL
    IF((0.5DO*GSU.GT.FF-FU).AND.(FF-FU.GT.GSU))
@ SL=DSL*0.5DO*GSU/(FU-FF+GSU)
    IF((GGG.GT.GSU).AND.(GSU.GT.2.DO*GGG))SL=DSL*GGG/(GSU-GGG)
    IF(SL.LT.1.05DO*DSL)SL=1.05DO*DSL
    IF(SL.GT.4.DO*DSL)SL=4.DO*DSL
    SL=SLU+SL
    FU=FF
    GSU=GGG
    GOTO 201
208 CONTINUE

C
C     TESTEN DES ABBRUCHKRITERIUMS
C     -----
C     ITNU=ITNU+1
C     KTR=0
C     CALL STOTES(N,XX,GG,P,FF,DF,SL,EPS,KTR,BG)
C     IF(KTR.EQ.1)RETURN

C
C     VORBEREITUNG FUER UPDATE-PROZEDUR
C     -----
C     YS=SL*(GGG-GS)
C     DO 401 I=1,N
C         Y(I)=GG(I)-G(I)
C         P(I)=SL*P(I)
401 CONTINUE
    IF((ITNU.EQ.1).AND.(N.GT.9))CALL HSCALE(N,H,Y,X,YS)

C
C     UPDATE-PROZEDUR
C     -----
C     CALL BFGS(N,H,P,Y,X,YS)

C
C     BESTIMMEN DER NEUEN SUCHRICHTUNG
C     -----
C     CALL HMULT(N,H,GG,P)
C     GS=0.DO
C     F=FF
C     DO 501 I=1,N
C         X(I)=XX(I)
C         G(I)=GG(I)
C         GS=GS+G(I)*P(I)
501 CONTINUE
    GOTO 101
END

```

```

C*****
C      SUBROUTINE  UOPTCO
C      -----
C      DIESE SUBROUTINE LOEST FREIE MINIMIERUNGSPROBLEME
C      VERWENDETE METHODE: QUASI-NEWTON MIT KONISCHEM MODELL
C      DIE WAHL VON AQER WIRD DURCH KTR GESTEUERT
C      VERSION: 16.08.84
C*****
C
C      SUBROUTINE UOPTCO(N,X,XX,G,GG,Y,P,H,EPS,FF,ITNU,IEV,KTR,BG)
C      IMPLICIT REAL*8 (A-H,O-Z)
C      DIMENSION X(N),XX(N),G(N),GG(N),Y(N),P(N),H(*)
C
C      VORBEREITUNG DES ERSTEN SCHRITTES
C      -----
C
C      IEVMAX=IEV
C      KTRCON=KTR
C      CALL FUN(N,X,F,G)
C      DF=DABS(F-FF)
C      IEV=1
C      ITNU=0
C      CALL HMULT(N,H,G,P)
C      GS=0.D0
C      BG=0.D0
C      DO 102 I=1,N
C         GS=GS+G(I)*P(I)
C         BG=DMAX1(BG,DABS(G(I)))
102  CONTINUE
C      KTR=1
C      IF (BG.LT.1.D-30)RETURN
101  CONTINUE
C      KTR=3
C      IF (GS.GE.0.D0)RETURN
C
C      LINE SEARCH
C      -----
C
C      SL=-2.D0*DF/GS
C      IF ((SL.GT.1.D0).OR.(ITNU.GT.N))SL=1.D0
C      SLU=0.D0
C      FU=F
C      GSU=GS
C      SLO=0.D0
201  DO 202 I=1,N
C         XX(I)=X(I)+SL*P(I)
202  CONTINUE
C      CALL FUN(N,XX,FF,GG)
C      IEV=IEV+1
C      KTR=2
C      IF (IEV.GE.IEVMAX)RETURN
C      GGS=0.D0
C      DO 203 I=1,N
C         GGS=GGS+GG(I)*P(I)
203  CONTINUE
C      DF=F-FF
C      IF (DF.LT.-1.D-2*GS*SL)GOTO 204
C      IF (GGS.LT.0.9D0*GS)GOTO205
C      ROQ=DF*KDF-GGS*GS*SL*SL
C      IF (ROQ.LE.DMIN1(1.D-3*GS*GS,1.D-30))GOTO 205
C      GOTO 208
204  SLO=SL
C      FO=FF
C      GSO=GGS
C      GOTO 206

```

```

205 IF(SLO.EQ.0.DO)GOTO 207
    SLU=SL
    FU=FF
    GSU=GGG
206 DSL=SLO-SLU
    Z=3.DO*(FU-F0)/DSL+GSU+GSO
    W=DSQRT(Z*Z-GSU*GSO)
    SL=DSL*(1.DO-(GSO+W-Z)/(GSO-GSU+2.DO*W))
    IF(SL.LT.0.05DO*DSL)SL=0.05DO*DSL
    IF(SL.GT.0.95DO*DSL)SL=0.95DO*DSL
    SL=SLU+SL
    GOTO 201
207 DSL=SL-SLU
    SLU=SL
    SL=2.DO*DSL
    IF((0.5DO*GSU.GT.FF-FU).AND.(FF-FU.GT.GSU))
    @ SL=DSL*0.5DO*GSU/(FU-FF+GSU)
    IF((GGG.GT.GSU).AND.(GSU.GT.2.DO*GGG))SL=DSL*GGG/(GSU-GGG)
    IF(SL.LT.1.05DO*DSL)SL=1.05DO*DSL
    IF(SL.GT.4.DO*DSL)SL=4.DO*DSL
    SL=SLU+SL
    FU=FF
    GSU=GGG
    GOTO 201
208 CONTINUE

```

```

C
C      TESTEN DES ABRUCHKRITERIUMS
C      -----
    ITNU=ITNU+1
    KTR=0
    CALL STOTES(N,XX,GG,P,FF,DF,SL,EPS,KTR,BG)
    IF(KTR.EQ.1)RETURN

```

```

C
C      VORBEREITUNG FUER UPDATE-PROZEDUR
C      -----
    GS=SL*GS
    GGS=SL*GGG
    RO=DSQRT(ROQ)
    GMA=-GS/(DF+RO)
    IF(DABS(GMA-1.DO).LT.1.D-5)GMA=1.DO
    YS=2.DO*RO

```

```

C
    IF (KTRCON.EQ.1) THEN
        BP2=0.DO
        DO 401 I=1,N
            BP2=BP2+P(I)*P(I)
401      CONTINUE
        FAK=(1.DO/GMA-1.DO)*GS/(BP2*SL)
        DO 402 I=1,N
            Y(I)=GG(I)*GMA-G(I)-FAK*P(I)
            P(I)=SL*P(I)*GMA*GMA
402      CONTINUE
    ENDIF

```

```

C
    IF (KTRCON.EQ.2) THEN
        DO 403 I=1,N
            P(I)=SL*P(I)*GMA*GMA
            Y(I)=GG(I)*GMA-G(I)/GMA
403      CONTINUE
    ENDIF

```

```

C
    IF (KTRCON.EQ.3) THEN
        FAKGG=GMA-(1.DO-GMA)*GS/(GMA*GGG-GS)

```

```

      FAKG=1. DO/GMA-(1. DO-GMA)*GGG/(GMA*GGG-GS)
      DO 404 I=1,N
        P(I)=SL*P(I)*GMA*GMA
        Y(I)=FAKGG*GG(I)-FAKG*G(I)
404    CONTINUE
      ENDIF
C
      IF (KTRCON.EQ. 4) THEN
        SS=0. DO
        DO 405 I=1,N
          SS=SS+P(I)*P(I)
405    CONTINUE
        FAK=(1. DO-GMA)*(GS/GMA+GGG)/(SS*SL)
        DO 406 I=1,N
          Y(I)=GG(I)-G(I)-FAK*P(I)
          P(I)=SL*P(I)*GMA*GMA
406    CONTINUE
      ENDIF
C
      IF (KTRCON.EQ. 5) THEN
        FAK=1. DO/GMA+(1. DO-GMA)*GGG/GS
        DO 407 I=1,N
          P(I)=SL*P(I)*GMA*GMA
          Y(I)=GG(I)-G(I)*FAK
407    CONTINUE
      ENDIF
C
      IF (KTRCON.EQ. 6) THEN
        FAK=(GMA*GMA*GGG-GS)/(GMA*(GGG-GS))
        DO 408 I=1,N
          P(I)=SL*P(I)*GMA*GMA
          Y(I)=FAK*(GG(I)-G(I))
408    CONTINUE
      ENDIF
C
      FAK=YS*GMA*GMA
      IF((ITNU.EQ. 1).AND.(N.GT. 9))CALL HSCALE(N,H,Y,X,FAK)
C
      UPDATE-PROZEDUR
C
      -----
      CALL BFGS(N,H,P,Y,X,FAK)
C
      BESTIMMEN DER NEUEN SUCHRICHTUNG
C
      -----
      CALL HMULT(N,H,GG,P)
      GGHGG=0. DO
      GHGG=0. DO
      F=FF
      DO 501 I=1,N
        GGHGG=GGHGG+GG(I)*P(I)
        GHGG=GHGG+G(I)*P(I)
        X(I)=XX(I)
        G(I)=GG(I)
501    CONTINUE
      FAK=(GMA*GGHGG-GHGG-2. DO*DF*GMA)/GS
      IF(FAK.LE. 1. D-50)FAK=1. DO
      DO 502 I=1,N
        P(I)=P(I)/FAK
502    CONTINUE
      GS=GGHGG/FAK
      GOTO 101
      END

```

```

C*****
C      SUBROUTINE  UOPTRA
C      -----
C      DIESE SUBROUTINE LOEST FREIE MINIMIERUNGSPROBLEME
C      VERWENDETE METHODE: QUASI-NEWTON MIT RATIONALEM MODELL
C      AGER // GQUER-G
C      VERSION: 16.08.84
C*****
C
C      SUBROUTINE UOPTRA(N,X,XX,G,G6,Y,P,H,EPS,FF,ITNU,IEV,KTR,BG)
C      IMPLICIT REAL*8 (A-H,O-Z)
C      DIMENSION X(N),XX(N),G(N),G6(N),Y(N),P(N),H(*)
C
C      VORBEREITUNG DES ERSTEN SCHRITTES
C      -----
C      IEVMAX=IEV
C      CALL FUN(N,X,F,G)
C      DF=DABS(F-FF)
C      IEV=1
C      ITNU=0
C      CALL HMULT(N,H,G,P)
C      GS=0. DO
C      BG=0. DO
C      DO 101 I=1,N
C         GS=GS+G(I)*P(I)
C         BG=DMAX1(BG,DABS(G(I)))
101  CONTINUE
C      KTR=1
C      IF(BG.LT.1.D-30)RETURN
102  CONTINUE
C      KTR=3
C      IF(GS.GE.0.DO)RETURN
C
C      LINE SEARCH
C      -----
C      SL=-2.DO*DF/GS
C      IF((SL.GT.1.DO).OR.(ITNU.GT.N))SL=1.DO
C      SLU=0.DO
C      FU=F
C      GSU=GS
C      SLO=0.DO
C      ISWIT=0
201  CONTINUE
C      DO 202 I=1,N
C         XX(I)=X(I)+SL*P(I)
202  CONTINUE
C      CALL FUN(N,XX,FF,G6)
C      IEV=IEV+1
C      KTR=2
C      IF(IEV.GE.IEVMAX)RETURN
C      GGS=0.DO
C      DO 203 I=1,N
C         GGS=GGS+G6(I)*P(I)
203  CONTINUE
C      DF=F-FF
C      IF(ISWIT.EQ.1)GOTO 208
C      IF(DF.GT.-GS*SL+1.D-30)THEN
C         ISWIT=1
C         GOTO 208
C      END IF

```

```

IF (DF, LT, -1, D-2*GS*SL) GOTO 204
IF (GGS, LT, 0, 9D0*GS) GOTO 205
ROQ=DF*DF-GGS*GS*SL*SL
IF (ROQ, LE, DMIN1(1, D-3*GS*GS, 1, D-30)) GOTO 205
GOTO 211
204 SLO=SL
FO=FF
GSO=GGS
GOTO 206
205 IF (SLO, EQ, 0, D0) GOTO 207
SLU=SL
FU=FF
GSU=GGS
206 DSL=SLO-SLU
Z=3, D0*(FU-FO)/DSL+GSU+GSO
W=DSQRT(Z*Z-GSU*GSO)
SL=DSL*(1, D0-(GSO+W-Z)/(GSO-GSU+2, D0*W))
IF (SL, LT, 0, 05D0*DSL) SL=0, 05D0*DSL
IF (SL, GT, 0, 95D0*DSL) SL=0, 95D0*DSL
SL=SLU+SL
GOTO 201
207 DSL=SL-SLU
SLU=SL
SL=2, D0*DSL
IF ((0, 5D0*GSU, GT, FF-FU), AND, (FF-FU, GT, GSU))
@ SL=DSL*0, 5D0*GSU/(FU-FF+GSU)
IF ((GGS, GT, GSU), AND, (GSU, GT, 2, D0*GGS)) SL=DSL*GGS/(GSU-GGS)
IF (SL, LT, 1, 05D0*DSL) SL=1, 05D0*DSL
IF (SL, GT, 4, D0*DSL) SL=4, D0*DSL
SL=SLU+SL
FU=FF
GSU=GGS
GOTO 201
208 CONTINUE
IF (DF, LT, -1, D-2*GS*SL) GOTO 209
IF (DF, GT, -GS*SL+1, D-30) GOTO 210
ROQ=DF*DF-GGS*GS*SL*SL
IF (ROQ, LE, 1, D-30) GOTO 210
GOTO 211
209 SLO=SL
SL=(SLU+SLO)/2, D0
GOTO 201
210 SLU=SL
IF (SLO, EQ, 0, D0) THEN
    SL=2, D0*SL
ELSE
    SL=(SLU+SLO)/2, D0
END IF
GOTO 201
211 CONTINUE

```

```

C
C      TESTEN DES ABRUCHKRITERIUMS
C      -----

```

```

ITNU=ITNU+1
KTR=0
CALL STOTES(N, XX, GG, P, FF, DF, SL, EPS, KTR, BG)
IF (KTR, EQ, 1) RETURN

```

```

C
C      VORBEREITUNG FUER UPDATE-PROZEDUR
C      -----
GS=SL*GS
GGS=SL*GGS
GMA=-(DF+GS)/(DF+GGS)
IF (DABS(GMA-1. DO). LT. 1. D-5) GMA=1. DO
YS=2. DO*ROQ/(-DF-GS)
FAK=(1. DO/GMA-1. DO)*DF/(GGS-GS)
FAKGG=1. DO+FAK
FAKG=1. DO/GMA+FAK
DO 301 I=1, N
    P(I)=SL*P(I)
    Y(I)=FAKGG*GG(I)-FAKG*G(I)
301 CONTINUE
IF ((ITNU.EQ. 1). AND. (N.GT. 9)) CALL HSCALE(N, H, Y, X, YS)
C
C      UPDATE-PROZEDUR
C      -----
CALL BFGS(N, H, P, Y, X, YS)
C
C      BESTIMMEN DER NEUEN SUCHRICHTUNG
C      -----
FAK=(1. DO/GMA-1. DO)/(GGS-GS)
DO 401 I=1, N
    Y(I)=FAK*(GG(I)-G(I))
401 CONTINUE
CALL HMULT(N, H, GG, P)
CALL HMULT(N, H, Y, G)
GGHGG=0. DO
AHGG=0. DO
AHA=0. DO
DO 402 I=1, N
    GGHGG=GGHGG-GG(I)*P(I)
    AHGG=AHGG-Y(I)*P(I)
    AHA=AHA-Y(I)*G(I)
402 CONTINUE
FAK=1. DO+AHGG
FAKMUE=GGHGG/(FAK+DSQRT(FAK*FAK-AHA*GGHGG))
F=FF
DO 403 I=1, N
    P(I)=P(I)-FAKMUE*G(I)
    X(I)=XX(I)
    G(I)=GG(I)
403 CONTINUE
GS=-GGHGG+FAKMUE*AHGG
GOTO 102
END

```

```

C*****
C
C  MATRIXMULTIPLIKATION
C  BERECHNET DAS PRODUKT DER MATRIX H MIT DEM VEKTOR G UND SPEICHERT
C  DEN NEGATIVEN RESULTATVEKTOR IN P
C
C*****

```

```

C
C  SUBROUTINE HMULT(N,H,G,P)
C  IMPLICIT REAL*8 (A-H,O-Z)
C  DIMENSION H(1),G(1),P(1)
C  DO 1 I=1,N
C    P(I)=0
C    DO 2 J=1,I
C      P(I)=P(I)-H(I*(I-1)/2+J)*G(J)
C  2  CONTINUE
C    IND=I+1
C    IF(IND.GT.N)GOTO 1
C    DO 3 J=IND,N
C      P(I)=P(I)-H(J*(J-1)/2+I)*G(J)
C  3  CONTINUE
C  1  CONTINUE
C  RETURN
C  END

```

```

C*****
C
C  MATRIXUPDATE
C  BILDET DIE NEUE MATRIX H-QUER NACH DER BFGS-FORMEL
C  UNTER VERWENDUNG DER ALTEN MATRIX H UND DER VEKTOREN P UND Y
C
C*****

```

```

C
C  SUBROUTINE BFGS(N,H,P,Y,G,YS)
C  IMPLICIT REAL*8 (A-H,O-Z)
C  DIMENSION H(1),P(1),Y(1),G(1)
C  CALL HMULT(N,H,Y,G)
C  YHY=0.DO
C  DO 1 I=1,N
C    YHY=YHY-G(I)*Y(I)
C  1  CONTINUE
C  FAK=1.DO+YHY/YS
C  DO 2 I=1,N
C    IND=I*(I-1)/2
C    DO 3 J=1,I
C      H(IND+J)=H(IND+J)+(FAK*P(I)*P(J)+P(I)*G(J)+P(J)*G(I))/YS
C  3  CONTINUE
C  2  CONTINUE
C  RETURN
C  END

```

```
C*****
C
C  SKALIERUNG DER ERSTEN MATRIX
C  METHODE NACH SHANNO-PHUA
C
```

```
C*****
C
```

```
      SUBROUTINE HSCALE(N,H,Y,X,YS)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION X(N),Y(N),H(*)
      CALL HMULT(N,H,Y,X)
      YHY=0.DO
      DO 1 I=1,N
        YHY=YHY-X(I)*Y(I)
1  CONTINUE
      FAK=YS/YHY
      DO 10 I=1,N*(N+1)/2
        H(I)=FAK*H(I)
10  CONTINUE
      RETURN
      END
```

```
C*****
C
```

```
C  STOPTEST
C  ASSESSMENTKRITERIUM NACH GILL-MURRAY
C
```

```
C*****
C
```

```
      SUBROUTINE STOTES(N,XX,GG,P,FF,DF,SL,EPS,KTR,BGG)
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION XX(N),GG(N),P(N)
      BGG=0.DO
      DO 1 I=1,N
        BGG=DMAX1(BGG,DABS(GG(I)))
1  CONTINUE
      IF(BGG.LT.1.D-30)GOTO 2
      IF(FF.GE.EPS)GOTO 3
2  CONTINUE
      KTR=1
3  CONTINUE
      RETURN
      END
```

4.3 Testmethode

Bei der nichtlinearen Optimierung mit Restriktionen wurde von Schittkowski [80] ein gutes Instrument zum Vergleich von Algorithmen entwickelt. Leider ist es unmöglich diese Technik der freien Optimierung anzupassen. Mit Problemen unseres Typs befasst sich eine Arbeit von Moré, Garbow und Hillstom [81]. Neben einem Katalog von 35 Testfunktionen wird darin eine komplette Testsoftware beschrieben, welche bei Verfahren für nichtlineare Gleichungssysteme, Nonlinear-least-squares und freie Minimierung verwendet werden kann. Für das Testen von Optimierungsverfahren werden nur 18 Beispielfunktionen zugezogen. Mit diesen 18 Funktionen wollen auch wir arbeiten und zählen sie deshalb mit ihren Namen auf:

1. Powell badly scaled (2)
2. Brown badly scaled (2)
3. Beal (2)
4. Helical valley (3)
5. Gauss (3)
6. Gulf research and development (3)
7. Box three-dimensional (3)
8. Wood (4)
9. Brown and Dennis (4)
10. Biggs EXP6 (6)
11. Watson (6, 9)
12. Rosenbrock extended (12, 24)
13. Powell singular extended (12, 24)
14. Penalty I (10, 20)
15. Penalty II (10, 20)
16. Variably dimensioned (10, 40)
17. Trigonometric (10, 40)
18. Chebyquad (8, 9)

In der Klammer hinter dem Namen ist die Dimension angegeben. Die letzten acht Funktionen sind von variabler Dimension, wir haben dafür je zwei konkrete Werte gewählt, sodass sich ein Total von 26 Funktionen ergibt.

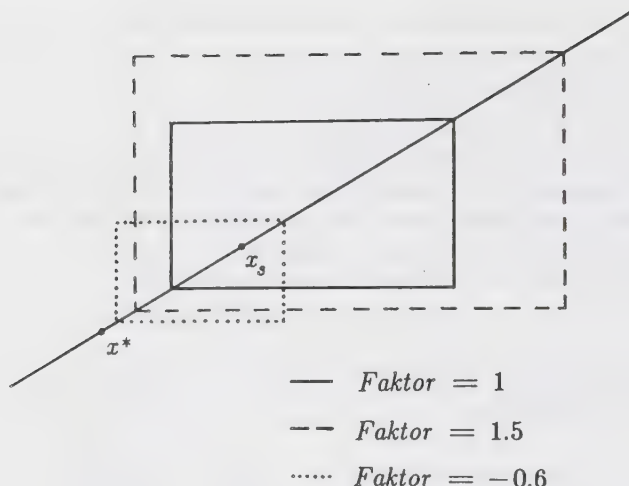
Das Verhalten der Algorithmen wird durch die Wahl des Startpunktes entscheidend beeinflusst. Als Beispiel haben wir in der folgenden Tabelle für die erweiterte Rosenbrock-Funktion mit Dimension 4 die benötigten Funktionsaufrufe notiert und zwar für drei Verfahren und drei Startpunkte.

Näherungsfunktion Startpunkt	quadratisch	konisch6	rational
(-4, 16, -4, 16)	102	116	159
(-16, 4, -16, 4)	230	201	254
(8, -1, -3, 1)	159	128	95

Um dieser Problematik gerecht zu werden, verwenden Moré, Garbow und Hillstom [81] neben dem Standardstartpunkt x_s , sofern möglich die beiden Ausgangspunkte $10 \cdot x_s$ und $100 \cdot x_s$. Gemeint wird damit, dass jede Komponente von x_s mit dem entsprechenden Faktor multipliziert wird. Diese Technik halten wir für ungenügend und verwenden stattdessen zufällig erzeugte Startpunkte ähnlich wie Hillstom [77]. Bei jeder Testfunktion ist ein Standardstartpunkt x_s und ein Standardminimum x^* bekannt. Daraus konstruieren wir komponentenweise unseren Zufallsstartpunkt x_r :

$$(x_r)_i = (x_s)_i + \text{Faktor} \cdot (x_s - x^*)_i \cdot (2r_i - 0.5). \quad (4.3.1)$$

Mit r_i werden hier (0, 1)-gleichverteilte Zufallszahlen bezeichnet. Der Startpunkt x_r wird somit zufällig aus einem quaderförmigen Bereich ausgewählt, welcher immer auch den Punkt x_s enthält. Wegen Overflowproblemen oder nahe zusammenliegender lokaler Minima ist es unmöglich die Quadergrösse bei allen Testfunktionen gleich zu wählen. Mit dem Parameter *Faktor* kann deshalb die Grösse des Quaders gesteuert werden. In der nachfolgenden Figur sind für einige Parameterwerte die zugehörigen Bereiche skizziert.



Bei unseren Testläufen haben wir für jede der 26 Beispielfunktionen zehn zufällige Startpunkte gewählt, sodass jedes der acht Verfahren auf 260 Probleme angesetzt wurde.

4.4 Resultate

Über die zehn Startpunkte wurde das arithmetische Mittel von der Anzahl der Iterationen (AMIT, Kolonne 1) sowie von der Zahl der Funktionsaufrufe (AMFA, Kolonne 3) berechnet. Die zugehörigen Varianzen sind in Kolonne 2 und 4 (VAIT, VAFA) notiert. Diese Angaben, wie auch das geometrische Mittel der Gradientennormen (GMGN, Kolonne 5), wurden nur bei jenen Testläufen ermittelt, wo die Genauigkeit gemäss Kriterium (4.2.2) erreicht wurde. In allen andern Fällen musste das Verfahren aufgrund einer anderen Bedingung gestoppt werden. Der Prozentsatz dieser Fehlschläge wird in Kolonne 6 angegeben und in Kolonne 7 (GMGN) schliesslich folgt das geometrische Mittel der Gradientennormen dieser Fehlschläge.

```

*****
NR. 1 * POWELL BADLY SCALED *
*****
          DIMENSION: 2          FAKTOR: -0.500D+00
          *****
          AMIT          VAFA          GMGN          %
          139.1      0.634D+02      171.7      0.589D-03      10.0
          122.3      0.118D+03      162.3      0.445D-03      10.0
          127.1      0.954D+02      166.9      0.157D+03      10.0
          119.1      0.649D+02      156.4      0.134D+03      10.0
          109.3      0.102D+03      154.3      0.192D+03      10.0
          113.9      0.384D+02      156.4      0.750D+03      10.0
          108.4      0.660D+02      156.7      0.182D+03      10.0
          120.0      0.443D+02      160.3      0.973D+02      10.0
          GMGN
          0.187D-13
          0.133D-06
          0.140D-09
          0.380D-06
          0.196D-07
          0.350D-06
          0.127D-06
          0.253D-06

```

```

*****
NR. 2 * BROWN BADLY SCALED *
*****
          DIMENSION: 2          FAKTOR: -0.100D+03
          *****
          AMIT          VAFA          GMGN          %
          17.8      0.193D+01      30.4      0.598D+01      20.0
          14.0      0.857D+01      28.1      0.755D+01      20.0
          16.9      0.461D+01      30.1      0.124D+02      20.0
          16.3      0.450D+01      29.8      0.116D+02      20.0
          16.9      0.327D+01      30.1      0.107D+02      20.0
          16.3      0.450D+01      29.8      0.116D+02      20.0
          12.4      0.313D+01      26.9      0.446D-03      20.0
          15.5      0.400D+01      29.1      0.698D+01      20.0
          0.242D-02
          0.158D-02
          0.604D-03
          0.450D-03
          0.173D-02
          0.446D-03
          0.129D-03
          0.208D-02
          GMGN
          0.477D+15
          0.483D+15
          0.274D+15
          0.214D+15
          0.665D+15
          0.221D+15
          0.795D+13
          0.180D+15

```

```

*****
NR. 3 * BEALE *
*****
          DIMENSION: 2          FAKTOR: -0.200D+01
          *****
          AMIT          VAFA          GMGN          %
          13.3      0.305D+02      16.3      0.560D+02      10.0
          11.3      0.103D+02      15.1      0.349D+02      10.0
          12.1      0.144D+02      15.6      0.330D+02      10.0
          11.3      0.925D+01      15.0      0.315D+02      10.0
          11.7      0.183D+02      15.1      0.431D+02      10.0
          11.6      0.140D+02      15.0      0.350D+02      10.0
          11.6      0.168D+02      15.0      0.433D+02      10.0
          11.8      0.132D+02      15.6      0.540D+02      10.0
          0.114D-07
          0.265D-08
          0.944D-08
          0.417D-08
          0.798D-08
          0.126D-08
          0.218D-07
          0.627D-08
          GMGN
          0.107D+03
          0.758D+02
          0.393D+02
          0.555D+02
          0.921D+02
          0.111D+03
          0.124D+03
          0.765D+02

```

 NR. 4 * HELICAL VALLEY * DIMENSION: 3 FAKTOR: 0.100D+01 *****

 AMIT VAIT AMFA VAFA GMGN % GMGN
 QUAD 29.1 0.788D+01 35.0 0.122D+02 0.594D-07 0.0 0.0
 CON1 29.2 0.907D+01 35.1 0.112D+02 0.216D-07 0.0 0.0
 CON2 29.2 0.796D+01 35.2 0.102D+02 0.445D-07 0.0 0.0
 CON3 27.3 0.734D+01 34.6 0.982D+01 0.550D-07 0.0 0.0
 CON4 27.7 0.846D+01 34.3 0.162D+02 0.842D-07 0.0 0.0
 CON5 28.0 0.109D+02 35.9 0.214D+02 0.456D-07 0.0 0.0
 CON6 27.9 0.101D+02 32.9 0.194D+02 0.460D-07 0.0 0.0
 RAT 28.3 0.490D+01 33.4 0.782D+01 0.598D-07 0.0 0.0

 NR. 5 * GAUSS * DIMENSION: 3 FAKTOR: 0.200D+04 *****

 AMIT VAIT AMFA VAFA GMGN % GMGN
 QUAD 10.4 0.782D+01 11.7 0.100D+02 0.225D-07 0.0 0.0
 CON1 9.1 0.143D+01 11.2 0.396D+01 0.448D-07 0.0 0.0
 CON2 9.0 0.156D+01 11.1 0.388D+01 0.191D-07 0.0 0.0
 CON3 8.8 0.129D+01 10.9 0.277D+01 0.167D-07 0.0 0.0
 CON4 8.9 0.188D+01 11.0 0.378D+01 0.447D-07 0.0 0.0
 CON5 9.2 0.218D+01 11.0 0.333D+01 0.161D-07 0.0 0.0
 CON6 8.6 0.933D+00 10.7 0.290D+01 0.489D-07 0.0 0.0
 RAT 8.9 0.121D+01 11.0 0.333D+01 0.315D-07 0.0 0.0

 NR. 6 * GULF RESEARCH AND DEVELOPMENT * DIMENSION: 3 FAKTOR: -0.200D+00 *****

 AMIT VAIT AMFA VAFA GMGN % GMGN
 QUAD 34.7 0.868D+01 42.9 0.788D+01 0.235D-07 0.0 0.0
 CON1 39.9 0.224D+03 50.3 0.500D+03 0.149D-07 0.0 0.0
 CON2 34.9 0.588D+01 42.7 0.269D+02 0.317D-07 0.0 0.0
 CON3 34.9 0.810D+01 41.9 0.810D+01 0.216D-07 0.0 0.0
 CON4 36.6 0.109D+03 45.1 0.180D+03 0.356D-07 0.0 0.0
 CON5 34.8 0.973D+01 42.8 0.773D+01 0.167D-07 0.0 0.0
 CON6 34.4 0.916D+01 42.6 0.114D+02 0.299D-07 0.0 0.0
 RAT 34.9 0.254D+01 43.3 0.105D+02 0.140D-07 0.0 0.0


```

*****
* BIGGS EXP6 *
*****
NR. 10
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      Z      GMGN
QUAD      116.1      0.341D+04      134.4      0.564D+04      0.132D-07      0.0      0.0
CON1      87.7      0.282D+03      102.8      0.495D+03      0.223D-07      0.0      0.0
CON2      93.1      0.517D+03      107.9      0.898D+03      0.164D-07      0.0      0.0
CON3      85.9      0.284D+03      97.3      0.405D+03      0.219D-07      0.0      0.0
CON4      87.8      0.341D+03      105.4      0.781D+03      0.100D-07      0.0      0.0
CON5      87.6      0.300D+03      101.0      0.416D+03      0.261D-07      0.0      0.0
CON6      86.5      0.348D+03      99.9      0.539D+03      0.132D-07      0.0      0.0
RAT      102.6      0.108D+04      116.7      0.158D+04      0.145D-07      0.0      0.0

```

```

*****
* WATSON *
*****
NR. 11
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      Z      GMGN
QUAD      115.9      0.100D+03      134.5      0.123D+03      0.268D-03      0.0      0.0
CON1      124.9      0.848D+02      149.4      0.158D+03      0.310D-03      0.0      0.0
CON2      117.7      0.164D+03      143.4      0.264D+03      0.342D-03      0.0      0.0
CON3      110.4      0.127D+03      131.1      0.263D+03      0.151D-03      0.0      0.0
CON4      123.3      0.203D+03      150.8      0.362D+03      0.372D-03      0.0      0.0
CON5      107.5      0.601D+02      130.3      0.149D+03      0.232D-03      0.0      0.0
CON6      103.0      0.651D+02      124.2      0.822D+02      0.400D-03      0.0      0.0
RAT      111.0      0.640D+02      130.4      0.827D+02      0.260D-03      0.0      0.0

```

```

*****
* WATSON *
*****
NR. 11
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      Z      GMGN
QUAD      145.8      0.184D+03      164.4      0.323D+03      0.660D-05      0.0      0.0
CON1      169.4      0.176D+04      195.6      0.270D+04      0.419D-05      0.0      0.0
CON2      147.6      0.528D+03      171.6      0.102D+04      0.757D-05      0.0      0.0
CON3      136.9      0.184D+03      156.1      0.321D+03      0.832D-05      0.0      0.0
CON4      178.7      0.826D+03      210.7      0.157D+04      0.811D-05      0.0      0.0
CON5      140.2      0.660D+03      160.3      0.114D+04      0.568D-05      0.0      0.0
CON6      133.0      0.346D+03      154.3      0.580D+03      0.419D-05      0.0      0.0
RAT      148.8      0.957D+03      168.5      0.142D+04      0.568D-05      0.0      0.0

```

* ROSENBROCK EXTENDED *										
NR. 12	*****				DIMENSION: 12		FAKTOR: 0.300D+01		*****	

	AMIT	VAIT	AMFA	VAFA	GMGN			z	GMGN	
QUAD	225.8	0.587D+03	241.9	0.780D+03	0.160D-06			0.0	0.0	
CON1	195.9	0.609D+03	209.4	0.675D+03	0.101D-06			0.0	0.0	
CON2	200.3	0.794D+03	212.0	0.892D+03	0.127D-06			0.0	0.0	
CON3	202.2	0.564D+03	215.8	0.734D+03	0.194D-06			0.0	0.0	
CON4	194.9	0.353D+03	209.3	0.424D+03	0.118D-06			0.0	0.0	
CON5	193.2	0.325D+03	209.0	0.389D+03	0.179D-06			0.0	0.0	
CON6	197.4	0.320D+03	210.7	0.404D+03	0.134D-06			0.0	0.0	
RAT	214.3	0.642D+03	221.2	0.832D+03	0.160D-06			0.0	0.0	

NR. 12	***** * ROSENBROCK EXTENDED *				***** DIMENSION: 24		***** FAKTOR: 0.200D+00	
	AMIT	VAIT	AMFA	VAFa	GMGN	%	GMGN	
QUAD	148.0	0.539D+03	157.5	0.634D+03	0.191D-06	0.0	0.0	
CGN1	105.0	0.106D+03	111.7	0.117D+03	0.106D-06	0.0	0.0	
CGN2	105.9	0.187D+03	112.9	0.212D+03	0.108D-06	0.0	0.0	
CGN3	100.5	0.108D+03	107.4	0.111D+03	0.688D-07	0.0	0.0	
CGN4	101.8	0.891D+02	108.6	0.965D+02	0.796D-07	0.0	0.0	
CGN5	104.0	0.998D+02	110.5	0.983D+02	0.951D-07	0.0	0.0	
CGN6	103.3	0.143D+03	108.1	0.154D+03	0.729D-07	0.0	0.0	
RAT	109.7	0.345D+03	115.1	0.377D+03	0.713D-07	0.0	0.0	

* POWELL SINGULAR EXTENDED *							

NR. 13	DIMENSION: 12						FAKTOR: 0.300D+01

	AMIT	VAIT	AMFA	VAF A	GMGN	%	GMGN
QUAD	227.8	0.480D+03	231.9	0.491D+03	0.260D-08	0.0	0.0
CON1	203.7	0.376D+03	214.2	0.416D+03	0.304D-07	0.0	0.0
CON2	210.6	0.295D+03	222.6	0.299D+03	0.381D-07	0.0	0.0
CON3	204.4	0.871D+03	214.7	0.103D+04	0.742D-07	0.0	0.0
CON4	200.6	0.277D+03	209.7	0.285D+03	0.419D-07	0.0	0.0
CON5	210.9	0.261D+03	223.5	0.329D+03	0.368D-07	0.0	0.0
CON6	199.9	0.401D+03	208.4	0.399D+03	0.620D-07	0.0	0.0
CAT	207.8	0.399D+03	216.0	0.413D+03	0.271D-07	0.0	0.0

```

*****
* POWELL SINGULAR EXTENDED *
*****
NR. 13
*****
          AMIT      VAIT      AMFA      VAFA      DIMENSION: 24      FAKTOR: 0.200D+00
          352.3      0.458D+03      355.1      0.497D+03      0.668D-08      0.0
          315.5      0.658D+03      321.8      0.724D+03      0.533D-08      0.0
          323.8      0.612D+03      331.3      0.602D+03      0.890D-08      0.0
          302.0      0.590D+03      307.8      0.672D+03      0.658D-08      0.0
          309.6      0.863D+03      318.1      0.957D+03      0.230D-08      0.0
          306.0      0.983D+03      314.4      0.109D+04      0.938D-08      0.0
          311.4      0.792D+03      317.0      0.831D+03      0.819D-08      0.0
          306.3      0.316D+03      311.3      0.346D+03      0.676D-08      0.0
          RAT
    
```

```

*****
* PENALTY I *
*****
NR. 14
*****
          AMIT      VAIT      AMFA      VAFA      DIMENSION: 10      FAKTOR: 0.100D+03
          101.6      0.212D+02      122.0      0.422D+02      0.421D-05      0.0
          76.5      0.221D+02      100.5      0.416D+02      0.414D-05      0.0
          79.2      0.353D+02      100.9      0.474D+02      0.304D-05      0.0
          75.7      0.336D+02      97.5      0.301D+02      0.260D-05      0.0
          70.6      0.267D+02      95.4      0.187D+02      0.604D-05      0.0
          76.3      0.185D+02      97.7      0.456D+02      0.608D-05      0.0
          70.8      0.282D+02      97.0      0.440D+02      0.579D-05      0.0
          79.5      0.469D+02      101.7      0.673D+02      0.429D-05      0.0
          RAT
    
```

```

*****
* PENALTY I *
*****
NR. 14
*****
          AMIT      VAIT      AMFA      VAFA      DIMENSION: 20      FAKTOR: 0.100D+02
          88.8      0.996D+01      106.0      0.229D+02      0.323D-05      0.0
          68.1      0.350D+02      89.5      0.227D+02      0.316D-05      0.0
          70.9      0.357D+02      90.9      0.392D+02      0.410D-05      0.0
          67.3      0.271D+02      86.4      0.238D+02      0.619D-05      0.0
          65.0      0.211D+02      88.0      0.404D+02      0.331D-05      0.0
          66.3      0.345D+02      85.9      0.299D+02      0.304D-05      0.0
          65.2      0.166D+02      87.0      0.196D+02      0.284D-05      0.0
          67.3      0.105D+02      89.2      0.200D+02      0.370D-05      0.0
          RAT
    
```

```

*****
* PENALTY II *
*****
NR. 15
*****
DIMENSION: 10
*****
FAKTOR: 0.200D+01
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      %
QUAD      367.1      0.113D+05      446.6      0.179D+05      0.475D-04      0.0
CON1      315.4      0.628D+04      413.8      0.122D+05      0.341D-04      0.0
CON2      337.7      0.838D+04      427.8      0.158D+05      0.395D-04      0.0
CON3      311.8      0.489D+04      386.2      0.815D+04      0.329D-04      0.0
CON4      312.7      0.565D+04      417.5      0.148D+05      0.319D-04      0.0
CON5      321.5      0.341D+04      397.5      0.603D+04      0.504D-04      0.0
CON6      310.3      0.544D+04      399.8      0.960D+04      0.336D-04      0.0
RAT      352.0      0.898D+04      422.3      0.137D+05      0.384D-04      0.0
  
```

```

*****
* PENALTY II *
*****
NR. 15
*****
DIMENSION: 20
*****
FAKTOR: 0.400D+00
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      %
QUAD      374.7      0.166D+04      425.9      0.271D+04      0.303D-03      0.0
CON1      341.6      0.245D+04      426.1      0.580D+04      0.189D-03      0.0
CON2      326.6      0.261D+04      399.8      0.506D+04      0.163D-03      0.0
CON3      311.8      0.137D+04      372.1      0.245D+04      0.286D-03      0.0
CON4      325.0      0.122D+04      406.0      0.229D+04      0.202D-03      0.0
CON5      317.7      0.159D+04      384.3      0.318D+04      0.126D-03      0.0
CON6      313.0      0.131D+04      377.2      0.218D+04      0.258D-03      0.0
RAT      327.1      0.212D+04      388.8      0.402D+04      0.313D-03      0.0
  
```

```

*****
* VARIABLY DIMENSIONED *
*****
NR. 16
*****
DIMENSION: 10
*****
FAKTOR: 0.300D+02
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      %
QUAD      55.2      0.929D+01      56.5      0.783D+01      0.706D-07      0.0
CON1      42.9      0.188D+02      56.5      0.343D+02      0.117D-07      0.0
CON2      39.2      0.251D+02      48.7      0.153D+02      0.397D-07      0.0
CON3      38.7      0.111D+02      42.8      0.973D+01      0.494D-07      0.0
CON4      42.1      0.201D+02      47.6      0.849D+01      0.222D-07      0.0
CON5      38.7      0.113D+02      44.8      0.862D+01      0.256D-07      0.0
CON6      39.9      0.157D+02      42.5      0.450D+01      0.215D-07      0.0
RAT      39.4      0.293D+01      42.1      0.143D+01      0.556D-07      0.0
  
```

```

*****
* VARIABLY DIMENSIONED *
*****
NR. 16
*****
          DIMENSION: 40
          FAKTOR: 0.100D+02
          *****
          AMIT      VAIT      AMFA      VAFA      GMGN      %
          QUAD      75.6      0.182D+03      78.9      0.190D+03      0.177D-06      0.0
          CON1      60.4      0.149D+02      69.7      0.309D+02      0.588D-07      0.0
          CON2      57.6      0.136D+02      66.9      0.210D+02      0.961D-07      0.0
          CON3      56.9      0.921D+01      61.1      0.121D+01      0.715D-07      0.0
          CON4      60.2      0.840D+01      65.8      0.507D+01      0.554D-07      0.0
          CON5      57.5      0.161D+02      64.1      0.150D+02      0.748D-07      0.0
          CON6      59.0      0.578D+01      62.9      0.121D+01      0.720D-07      0.0
          RAT       58.6      0.338D+01      62.5      0.117D+01      0.495D-07      0.0
          GMGN
  
```

```

*****
* TRIGONOMETRIC *
*****
NR. 17
*****
          DIMENSION: 10
          FAKTOR: -0.200D+01
          *****
          AMIT      VAIT      AMFA      VAFA      GMGN      %
          QUAD      31.1      0.126D+03      33.0      0.156D+03      0.142D-07      10.0
          CON1      22.0      0.595D+02      23.8      0.712D+02      0.120D-07      10.0
          CON2      21.7      0.548D+02      23.7      0.645D+02      0.159D-07      10.0
          CON3      21.8      0.624D+02      23.3      0.678D+02      0.175D-07      10.0
          CON4      22.0      0.713D+02      23.6      0.783D+02      0.116D-07      10.0
          CON5      22.0      0.695D+02      23.8      0.794D+02      0.149D-07      10.0
          CON6      22.3      0.685D+02      24.0      0.880D+02      0.202D-07      10.0
          RAT       22.8      0.607D+02      24.3      0.713D+02      0.173D-07      10.0
          GMGN
  
```

```

*****
* TRIGONOMETRIC *
*****
NR. 17
*****
          DIMENSION: 40
          FAKTOR: -0.200D+01
          *****
          AMIT      VAIT      AMFA      VAFA      GMGN      %
          QUAD      43.3      0.675D+02      44.6      0.713D+02      0.192D-07      10.0
          CON1      27.9      0.191D+02      29.2      0.189D+02      0.131D-07      10.0
          CON2      28.0      0.190D+02      29.3      0.190D+02      0.173D-07      10.0
          CON3      27.4      0.243D+02      29.0      0.218D+02      0.148D-07      10.0
          CON4      27.8      0.217D+02      28.9      0.196D+02      0.133D-07      10.0
          CON5      28.0      0.233D+02      29.1      0.211D+02      0.121D-07      10.0
          CON6      28.3      0.235D+02      29.3      0.235D+02      0.137D-07      10.0
          RAT       28.6      0.213D+02      29.6      0.213D+02      0.167D-07      10.0
          GMGN
  
```

```

*****
* CHEBYQUAD *
*****
NR. 18          DIMENSION: 8          FAKTOR: 0.600D+01
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      %      GMGN
QUAD      37.2      0.584D+02      46.6      0.947D+02      0.450D-03      0.0      0.0
CUN1      36.3      0.471D+02      47.8      0.145D+03      0.371D-03      0.0      0.0
CON2      38.0      0.633D+02      48.3      0.104D+03      0.471D-03      0.0      0.0
CON3      38.6      0.754D+02      48.3      0.135D+03      0.467D-03      0.0      0.0
CON4      39.5      0.130D+03      49.0      0.138D+03      0.319D-03      0.0      0.0
CON5      40.2      0.786D+02      51.2      0.138D+03      0.298D-03      0.0      0.0
CON6      37.9      0.885D+02      46.8      0.123D+03      0.444D-03      0.0      0.0
RAT       42.8      0.104D+03      52.1      0.185D+03      0.375D-03      0.0      0.0

```

```

*****
* CHEBYQUAD *
*****
NR. 18          DIMENSION: 9          FAKTOR: 0.600D+01
*****
      AMIT      VAIT      AMFA      VAFA      GMGN      %      GMGN
QUAD      47.3      0.687D+02      56.3      0.109D+03      0.473D-07      0.0      0.0
CON1      48.6      0.336D+02      55.5      0.492D+02      0.464D-07      0.0      0.0
CON2      52.3      0.656D+02      62.2      0.133D+03      0.426D-07      0.0      0.0
CON3      49.4      0.179D+03      55.9      0.237D+03      0.420D-07      0.0      0.0
CON4      47.0      0.480D+02      53.2      0.533D+02      0.311D-07      0.0      0.0
CON5      49.1      0.623D+02      56.6      0.116D+03      0.313D-07      0.0      0.0
CON6      50.1      0.919D+02      58.1      0.119D+03      0.485D-07      0.0      0.0
RAT       47.9      0.965D+02      54.0      0.118D+03      0.583D-07      0.0      0.0

```

Bei unseren Testläufen war es nur für Funktion 7 nötig den Parameter *Faktor* abnormal klein zu wählen. Dies liegt offensichtlich an der miserablen Qualität des zugehörigen Standardstartpunktes. Interessanterweise liefert das Verfahren mit rationaler Näherungsfunktion hier Spitzenergebnisse.

Stellt man die gewöhnliche Variable-Metrik-Methode (*quad*) allen andern Verfahren (*con1* bis *con6* und *rat*) gegenüber, so sind drei Fälle zu unterscheiden:

- Die gesamte Gruppe der nichtquadratischen Verfahren erbringt bessere Resultate.
(Testfunktionen: 1, 3, 5, 9, 10, 12, 13, 14, 15, 16, 17).
- Die einzelnen nichtquadratischen Verfahren verhalten sich unterschiedlich.
(Testfunktionen: 2, 7, 8, 11)
- Die nichtquadratischen Verfahren liefern keine bessern aber auch keine wesentlich schlechteren Ergebnisse.
(Testfunktionen: 4, 6, 18)

Vergleicht man die verschiedenen nichtquadratischen Verfahren untereinander, so schneiden unseres Erachtens die Varianten *con3* und *con6* insgesamt am besten ab. Obwohl die Unterschiede klein ausfallen, erfüllen sich damit die Erwartungen, welche in Kapitel 2.2 an die Wahl der Norm geknüpft waren. Diese beiden Varianten zeigen bei 14 der 18 Testfunktionen bessere Ergebnisse als die gewöhnliche Variable-Metrik-Methode.

Die bisherigen experimentellen Untersuchungen über Verfahren mit nichtquadratischen Näherungsfunktionen sind wenig aufschlussreich. Sorensen [80] gibt die Ergebnisse von sechs Testfunktionen an, welche für sein Kollineares-Skalierungs-Verfahren insgesamt ein positives Bild zeichnen. Dagegen präsentiert Grandinetti [84] für dasselbe Verfahren die Daten von fünf Testfunktionen mit vorwiegend negativem Untersuchungsergebnissen. Beide Arbeiten vergleichen zwei völlig verschiedene Implementationen miteinander, die Aussagen sind deshalb nur von beschränkter Gültigkeit. Für ein Verfahren, das die Hesse-Matrix der Problemfunktion zusammen mit einer konischen Näherungsfunktion verwendet, rapportiert Schnabel [83] einige Prozentzahlen von umfangreichen Testversuchen, welche aber nur teilweise erfolgreich ausfielen.

Die Resultate unserer experimentellen Untersuchung zeigen, dass nichtquadratische Näherungsfunktionen bei Variable-Metrik-Verfahren bezüglich Anzahl Iterationen und Funktionsaufrufen eine Verbesserung bringen. Demgegenüber hält sich der zusätzliche algorithmische Aufwand, zumindest was die konischen Verfahren betrifft, in engen Grenzen.

Literaturliste

- J. Barzilai und A. Ben-Tal [1982], *Nonpolynomial and inverse interpolation for line search: synthesis and convergence rates*,
SIAM J. Numer. Anal. 19, 1263 - 1277
- K. W. Brodlie [1977], *An assessment of two approaches to variable metric methods*,
Math. Prog. 12, 344 - 355
- C. G. Broyden [1969], *A new double-rank minimization algorithm*,
AMS Notices 16, 670
- C. G. Broyden, J. E. Dennis und J. J. Moré [1973], *On the local and superlinear convergence of quasi-Newton methods*, J. Inst. Math. Appl. 12, 223 - 246
- W. C. Davidon [1959], *Variable metric method for minimization*, Argonne National Laboratory Report ANL-5990 (Rev.),
- W. C. Davidon [1975], *Optimally conditioned optimization algorithms without line searches*,
Math. Prog. 9, 1 - 30
- W. C. Davidon [1980], *Conic Approximations and collinear scalings for optimizers*,
SIAM J. Numer. Anal. 17, 268 - 281
- J. E. Dennis und J. J. Moré [1974], *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp. 28, 549 - 560
- J. E. Dennis und J. J. Moré [1977], *Quasi-Newton methods, motivation and theory*,
SIAM Review 19, 46 - 89
- J. E. Dennis und R. B. Schnabel [1979], *Least change secant updates for quasi-Newton methods*, SIAM Review 21, 443 - 459
- L. C. W. Dixon [1972], *Quasi-Newton algorithms generate identical points*,
Math. Prog. 2, 383 - 387, (Math. Prog. 3, 345 - 358)
- R. Fletcher [1970], *A new approach to variable metric algorithms*,
Comp. J. 13, 317 - 322

- R. Fletcher [1980], *Practical methods of optimization*, vol. 1: *unconstrained optimization*, John Wiley & Sons, New York
- R. Fletcher und M. J. D. Powell [1963], *A rapidly convergent descent method for minimization*, Comp. J. 6, 163 - 168
- P. E. Gill und W. Murray [1979], *Performance Evaluation for optimization software*, in *Performance evaluation of numerical software*, (L. Fosdick ed.), North-Holland, Amsterdam, 221 - 234
- D. Goldfarb [1970], *A family of variable-metric methods derived by variational means*, Math. Comp. 24, 23 - 26
- L. Grandinetti [1980], *Factorized variable metric algorithms for unconstrained optimization*, in *Optimization techniques*, part 2, (K. Iracki, K. Malanowski, S. Walukiewicz eds.), Springer-Verlag, Berlin, 52 - 61
- L. Grandinetti [1984], *Some investigations in a new algorithm for nonlinear optimization based on conic models of the objective function*, J. Optim. Theory Appl. 43, 1 - 21
- J. Greenstadt [1970], *Variations on variable-metric methods*, Math. Comp. 24, 1 - 18
- K. E. Hillstom [1977], *A simulation test approach to the evaluation of nonlinear optimization algorithms*, ACM Trans. Math. Softw. 3, 305 - 315
- J. J. Moré, B. S. Garbow und K. E. Hillstom [1981], *Testing unconstrained optimization software*, ACM Trans. Math. Softw. 7, 17 - 41
- J. M. Ortega und W. C. Rheinboldt [1970], *Iterative solution of nonlinear equations in several variables*, Academic Press, New York
- K. Schittkowski [1980], *Nonlinear programming codes; information, tests, performance*, Springer-Verlag, Berlin
- R. B. Schnabel [1983], *Conic methods for unconstrained minimization and tensor methods for nonlinear equations*, in *Mathematical programming, the state of the art*, (A. Bachem, M. Gröschel, B. Korte eds.), Springer-Verlag, Berlin, 417 - 438

- D. F. Shanno [1970], *Conditioning of quasi-Newton methods for function minimization*, Math. Comp. 24, 647 - 656
- D. F. Shanno und K. H. Phua [1978a], *Matrix conditioning and nonlinear Optimization*, Math. Prog. 14, 149 - 160
- D. F. Shanno und K. H. Phua [1978b], *Numerical comparison of several variable-metric algorithms*, J. Optim. Theory Appl. 25, 507 - 518
- D. C. Sorensen [1980], *The Q-superlinear convergence of collinear scaling algorithm for unconstrained optimization*, SIAM J. Numer. Anal. 17, 84 - 114

Inhaltsverzeichnis

1. EINLEITUNG	1
1.1 Überblick über die Variable-Metrik-Mehoden	1
1.2 Quadratische Näherungsfunktion	3
2. VERFAHREN MIT NICHTQUADRATISCHER NÄHERUNGSFUNKTION	5
2.1 Konische Näherungsfunktion	5
2.2 Rationale Näherungsfunktion	13
3. KONVERGENZEIGENSCHAFTEN	20
3.1 Lokale Konvergenz bei Variable-Metrik-Verfahren	20
3.2 Modifikationen im Beweis bei Verwendung konischer Näherungsfunktionen	24
3.3 Modifikationen im Beweis bei Verwendung rationaler Näherungsfunktionen	31
3.4 Superlineare Konvergenz	34
4. EMPIRISCHE UNTERSUCHUNG	36
4.1 Einleitung	36
4.2 Implementation der Verfahren	36
4.3 Testmethode	48
4.4 Resultate	50
Literaturliste	61

LEBENS LAUF von Fortunat Martin Schmid

Am 9. März 1951 wurde ich in Sax(SG) als Bürger der Stadt St. Gallen geboren. Nach dem Besuch der Primarschule in Winterthur und Bäretswil(ZH) trat ich 1964 ins Gymnasium in Wetzikon ein und wechselte 1969 an die Kantonschule Aarau, wo ich 1971 mit der Maturität des Typs C abschloss. Im selben Jahr immatrikulierte ich mich an der Universität Zürich mit dem Hauptfach Mathematik. 1977 bestand ich die Diplomprüfung. Seit 1974 bin ich Lehrer für Mathematik an der Kantonsschule Wiedikon in Zürich. 1980 absolvierte ich die Prüfung für das höhere Lehramt. Von 1979 bis 1983 war ich als Assistent am Institut für Operations Research tätig. Seit dem Herbst 1982 bin ich Lehrbeauftragter der Rechts- und Staatswissenschaftlichen Fakultät.

Verzeichnis der Dozenten, deren Vorlesungen ich besucht habe:

E. Brun, H. Bühlmann, P. Demuth, H. Garbers, H. Gross, K. Hässig, H. Jarchow, G. Karrer, M. Köhler, P. Kall, H.R. Schwarz, H. Storrer, K. Strebel, F. Ulmer, N. Wirth, O. Woodtli.

Die vorliegende Dissertation entstand unter der Leitung von Prof. Dr. P. Kall. Ihm möchte ich an dieser Stelle für seine Unterstützung meinen herzlichen Dank aussprechen.

